

VOL. 50, No. 5

SEPTEMBER, 1953

# Psychological Bulletin

## CONTENTS

### ARTICLES:

*A Brief History of Clinical Psychology:* ROBERT I. WATSON, 321  
*Psychology in Italy:* HENRYK MISIAK AND VIRGINIA M. STAUDT, 347  
*The Measurement of Individual Differences in Originality:* R. C. WILSON, J. P. GUILFORD AND P. R. CHRISTENSEN, 362  
*Correcting the Kuder-Richardson Reliability for Dispersion of Item Difficulties:* PAUL HORST, 371  
*Models for Testing the Significance of Combined Results:* LYLE V. JONES AND DONALD W. FISKE, 375  
*Historical Note on the Rating Scale:* DOUGLAS G. ELLSON AND ELIZABETH COX ELLSON, 383  
*A Brief Note on One-Tailed Tests:* C. J. BURKE, 384  
*A Note on the Recognition and Interpretation of Composite Factors:* WAYNE S. ZIMMERMAN, 387  
*A Rejoinder to Zimmerman's Note:* JOHN W. FRENCH, 390

### BOOK REVIEWS:

DE GRAZIA'S *Errors of psychotherapy* and LINDNER'S *Prescription for rebellion:* GEORGE W. ALBEE, 391  
BENEDEK'S *Studies in psychosomatic medicine: psychosexual functions in women:* M. ERIK WRIGHT, 392  
MAIER'S *Principles of human relations:* D. J. MOFFIE, 394  
BROWER AND ABT'S *Progress in clinical psychology:* JULIUS WISHNER, 395  
LANSING'S *Cowdry's problems of ageing:* WAYNE DENNIS, 396  
PIÉRON, PICHOT, FAVERGÉ AND STOETZEL'S *Méthodologie psychotechnique:* JOSEF BROŽEK, 397  
GILBERT'S *Understanding old age:* NATHAN W. SHOCK, 399  
REIK'S *The secret self:* DAN L. ADLER, 400

PUBLISHED BIMONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

WAYNE DENNIS, Editor

*Brooklyn College*

EDWARD GIRDEN, Associate Editor (Book Reviews)

*Brooklyn College*

ROBERT L. THORNDIKE, Associate Editor (Statistics)

*Teachers College, Columbia University*

LORRAINE BOUTHELET, Managing Editor

---

*The Psychological Bulletin* contains evaluative reviews of the literature in various fields of psychology, methodological articles, critical notes, and book reviews. This JOURNAL does not publish reports of original research or original theoretical articles.

*Editorial communications* and manuscripts should be sent to Wayne Dennis, Department of Psychology, Brooklyn College, Brooklyn 10, New York. Books for review should be sent to Edward Girden, at the same address.

*Preparation of articles for publication.* Authors are strongly advised to follow the general directions given in the "Publication Manual of the American Psychological Association" (*Psychological Bulletin*, 1952, 49 (No. 4, Part 2), 389-440). Special attention should be given to the section on the preparation of the references (pp. 432-440), since this is a particular source of difficulty in long reviews of research literature. *All copy must be double spaced, including the references.*

*Reprints.* Fifty free offprints are given to contributors of articles, notes, and special reviews. Five copies of the JOURNAL are supplied gratis to the authors of book reviews.

*Business communications*—including subscriptions, orders of back issues, and changes of address—should be sent to the American Psychological Association, 1333 Sixteenth Street N.W., Washington 6, D.C.

*Annual subscription: \$8.00 (Foreign \$8.50). Single copies, \$1.50.*

---

PUBLISHED BIMONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1333 Sixteenth Street N.W., Washington 6, D.C.

Published as second class mail matter at the post office at Washington, D.C., under the act of March 3, 1879, additional entry at the post office at Menasha, Wisconsin. Acceptance for mailing at special rate of postage under the provisions of Sec. 34-40 Part (D) provided for in Section 334, act of February 25, 1925, authorized August 6, 1947. Printed in U.S.A.

Copyright, 1953, by The American Psychological Association, Inc.

# Psychological Bulletin

## A BRIEF HISTORY OF CLINICAL PSYCHOLOGY

ROBERT I. WATSON  
*Northwestern University*<sup>1</sup>

Clinical psychologists have been surprisingly ahistorical. Very little thought has been given to, and less written about, the origin and development of clinical psychology. In the literature there are articles and books which interpret historically various special aspects or evaluate related fields, some of which have been of considerable help in preparing this paper. Nevertheless, whatever the reasons, there is no available general account of the history of clinical psychology from the perspective of today.

In part, this neglect is due to the upsurge of interest in clinical psychological activity during and following the second World War. Since then, clinical psychologists have had little time to spend inquiring into their origins. Then too, their day-to-day activities impress them as so new and vital that they are hardly to blame for tacitly accepting the belief that they are pioneers and that somewhere in the chaos of war and its aftermath was born a new profession having little or no relation to what went before. The state of affairs today is curiously reminiscent of the situation found by Kimball Young in 1923 in tracing the history of mental testing. He remarked, "Making history on

every hand as we are, we have a notion that we somehow have escaped history" (121, p. 1).

To capture in full measure the sweep and continuity of the history of clinical psychology is beyond the competence of the reviewer, to say nothing of space limitations. In order to do justice to all aspects of the subject one would have to deal with the complex history of the psychology of motivation and dynamic psychology. Similarly, all the ramifications of the relation of clinical psychology to the rest of the field of psychology, of which it is an integral part, as well as an account of the history of test development would have to be considered.

The present account, perchance, presents an examination of men and ideas influential in shaping clinical psychology. But, since psychology is now a profession, attention must also be devoted to those internal and external controls which characterize a profession and to the settings in which the professional practice is conducted.

In presenting a historical account the question arises concerning the most appropriate date at which to begin. With some justification it was decided that since clinical psychology, as we know it, arose at about the turn of the present century it would be appropriate to begin with the im-

<sup>1</sup> This article was written while the author was at the Washington University School of Medicine.

mediate forerunners of this first generation of clinical psychologists. The origins of clinical psychology, the first major section of this account, are to be found in the psychometric and dynamic traditions in psychology; the psychologist in the settings of the psychological clinic, child guidance, mental hospitals, institutions for the mentally defective; and the beginnings of psychology as a profession. Somewhat arbitrarily this early pioneer work is considered to come to a close with the end of the second decade of this century. This is followed by a section concerned with clinical psychology in the twenties and thirties. The same topics just mentioned, e.g., the dynamic tradition and psychology as a profession, are again considered. The work of psychologists in the armed services during the second World War and its effect upon psychology in the postwar period are next evaluated. A brief overview of clinical psychology today closes the account.

#### THE ORIGINS OF CLINICAL PSYCHOLOGY

##### *The Psychometric Tradition in Psychology*

This tradition, one of the headwaters from which clinical psychology sprang, was, in turn, a part of the scientific tradition of the nineteenth century. With all the limitations with which it is charged today, it is to this movement that the clinical psychologist owes much of his scientific standing and tradition. Whenever a clinical psychologist insists upon objectivity and the need for further research, he is, wittingly or otherwise, showing the influence of this tradition. Moving from Galton through Binet and Terman, this tradition met the demand that psychology, if it was to be-

come a science, must share with other sciences the respect for quantitative measurement.

Psychometrics as a tool for clinical psychology owes its beginnings to Francis Galton (53) in England. Grappling as he was with the problem of individual differences, he and his followers did much to lay the groundwork for the investigation of ability by using observations of an individual's performance for information on individual differences. He thus founded mental tests.

In 1890 Cattell (29) introduced the term "mental tests" in an article describing tests which he had used at the University of Pennsylvania. Even at that date he was pleading for standardization of procedure and the establishment of norms. From the time of his days as a student of Wundt's, Cattell was interested in the problem of individual differences and did much to stimulate further investigation. Along with Thorndike and Woodworth, he also stressed dealing with individual differences by means of statistical analysis—a really new approach at this time. Some of these investigations, both from Cattell's laboratory and from others in various parts of the country, made positive contributions to various facets of the problem of psychometric measurement. For example, Norsworthy (82) in 1906 compared normal and defective children by means of tests and found the latter not a "species apart," pointing out that the more intelligent of the feeble-minded were practically indistinguishable from the least intelligent of the normal.

Most of the investigations of the time were concerned with simple sensorimotor and associative functions and were based on the assumption that intelligence could be reduced to sensations and motor speed,

an attempt which, as is now known, was doomed to failure. Furthermore, although more suitable verbal material was used, the studies of college students at Cornell, such as Sharp's (95), and the Wissler study (113) at Columbia, were found to be essentially nonpredictive. What the workers failed to take into account was the fact that college students are a highly selected group having a considerably restricted range. The negative finding of these studies effectively blocked further investigation at the college level for years. When one stops to consider that the dominant systematic position of the day was the structuralism of Titchener, who had banished tests as nonscientific, it is no wonder that "tests" were viewed with at least a touch of condescension.

In the meantime Binet had been working in France developing his tests based on a wider sampling of behavior than had yet been used. His success in dealing with the intellectual classification of Paris school children is well known and needs no amplification at this point. The translation of his tests and their use in this country followed shortly after the turn of the century. It was Goddard (simultaneously with Healy), a student of G. Stanley Hall, who introduced the Binet tests to this country. Through a visit abroad and contact with Decroly, he became acquainted with Binet's work (121). In 1910 he began publishing findings with the test and in 1911 published his revision of the 1908 Binet Scale. This revision, along with Kuhlmann's, also published in 1911, gained some popularity among clinicians, but the subsequent development by Terman far overshadowed their work.

Probably the test that had the most influence upon trends in clinical psychology was the Terman Revision

of the Binet Scale (83). In fact, for years the major task of the clinical psychologist was to administer the Stanford-Binet. In view of the importance of this test it is desirable to present in some detail the background of its development.

Lewis M. Terman (101) received his graduate training at Clark just after the turn of the century under Hall, Sanford, and Burnham. As Terman put it, "For me, Clark University meant briefly three things: freedom to work as I pleased, unlimited library facilities and Hall's Monday evening seminar. Anyone of these outweighed all the lectures I attended" (101, p. 315). This influence of Hall's was more from the enthusiasm he inspired and the wide scope of his interests than from his scientific caution and objectivity. Sanford was his doctoral adviser, but Terman chose his own problem in the area of differentiation of "bright" and "dull" groups by means of tests and worked it through more or less independently.

By a severely limited survey such as this it would be easy to give the impression that little or nothing else was being done along the lines under discussion except that reported. Terman was not alone in his interest in the development and standardization of tests by any manner of means. In his autobiography Terman (101) mentions as known to him in 1904 the work of Binet, Galton, Bourdon, Oehrni, Ebbinghaus, Kraepelin, Aschaffenburg, Stern, Cattell, Wissler, Thorndike, Gilbert, Jastrow, Bolton, Thompson, Spearman, Sharp, and Kuhlmann.

At the suggestion of Huey, who had been working in Adolf Meyer's clinic at Johns Hopkins, Terman, undeterred by the prevailing hostile attitude of most psychologists, began

work with the 1908 Binet Scale and in 1916 published the Stanford Revision of the Binet-Simon tests. Terman's interest in both the test and results from it continued unabated, resulting in still another revision in 1937.

Performance tests, so necessary for work with the linguistically handicapped, actually antedated the Stanford-Binet. The Seguin, Witmer, and Healy form boards and other performance tests were already in clinical use. Norms, although not lacking, were undeveloped, and the directions placed a high premium on language. What seemed to be needed was a battery of performance tests sampling a variety of functions and not as dependent upon language. Among the earliest to appear and to come into fairly common use was the Pintner-Paterson Scale of Performance Tests (85), published in 1917. Included in this scale were several form boards, a manikin and a feature-profile construction test, a picture completion test, a substitution test, and a cube-tapping test.

Another major step was the development of group tests under the impetus of the need for large scale testing of recruits in World War I. This testing program is described with a wealth of detail by Yerkes (120). Although group tests were not unknown before the war, as witness those described in Whipple's *Manual of Mental and Physical Tests* (110), the need for quick appraisal of the basic intelligence of a large number of men provided the impetus for extensive development. The *Alpha* scale for literate English-speaking recruits and the *Beta* scale for illiterates and non-English-speaking recruits were developed rapidly under this demand. The Woodworth Personal Data Sheet (118), the first of a

long line of psychoneurotic inventories, also was a product of military needs. So successful were these tests in overcoming the prejudices against testing both within the field of psychology and in the general public that after the war a veritable flood of group tests appeared. Many extensive surveys in the public schools were made for classificatory purposes. Further developments in this tradition during the twenties and thirties will be appraised after examination of other aspects of the origin of clinical psychology.

#### *The Dynamic Tradition in Psychology*

A major source of influence contributing to the growth of clinical psychology was the thinking and writing of the "Boston group" who promulgated "the new psychology"—William James, G. Stanley Hall, and their associates. Although in no way could they be labeled clinical psychologists, their thinking was much closer to the heart of the clinical psychology movement and to progressive psychiatry than was the structural point of view of Titchener. Heresy though it may be, it cannot be denied that at that time academic psychology had relatively little to contribute to clinical psychology. Psychology, to be sure, had been placed by Fechner, Helmholtz, Wundt, Kraepelin, and others upon a scientific, quantitative foundation instead of being permitted to remain an indistinguishable cohort of philosophy. This was an essential step without which there could have been no clinical psychology; nevertheless, a sensationalistic approach to conscious intellectual experience offered relatively little for the clinical method and the profession with which it was to be associated.

The psychiatry of the day was in

the main concerned with pathology and the search for an explanation of mental disturbances in disease processes. Kraepelin (68) had introduced clarity through his classification of mental disease, but at the expense of deeper understanding. Based upon symptoms and primarily descriptive in character, his classification served to diminish—even to discourage—in its users any urge toward understanding of psychological dynamics.

French psychiatric thinking and research profoundly influenced James (80). The work of Janet and Charcot was particularly important in this connection. With Morton Prince, he did much to stimulate interest in the phenomena of dissociation, feeling as he did that it was a fruitful method of investigation of personality functioning. Early in his career he recognized the value of a clinical approach which led him "whenever possible to approach the mind by way of its pathology" (77, p. 20).

The influence of James was expressed primarily through his *Principles of Psychology* (65), published in 1890, and to a lesser degree by his *Varieties of Religious Experience* (66), published in 1902. Both of these works were sufficiently removed from the otherwise prevailing psychological thinking of his day to be considered major pre-Freudian, dynamic influences. The choice of the term "dynamic" in this context is neither idle nor wishful thinking. James himself used the term to distinguish his point of view from the structural approach of Titchener (86).

Concerning the influence of the *Principles*, Morris had this to say:

Great books are either reservoirs or watersheds. They sum up and transmit the antecedent past, or they initiate the flow of the future. Sixty years after its publication, the *Principles* appears to be

one of the major watersheds of twentieth-century thought. Directly or indirectly, its influence had penetrated politics, jurisprudence, sociology, education and the arts. In the domain of psychology, it had foreshadowed nearly all subsequent developments of primary importance. Viewed retrospectively, the permanent significance of the *Principles* was incentive. It explored possibilities and indicated directions. These led, eventually, into social, applied and experimental psychology; into the study of exceptional mental states, subliminal consciousness and psychopathology. Because of its extreme fertility in the materials for hypothesis, most of the competitive schools of psychological theory that arose during the first half of the century could claim common ancestry in the *Principles* for at some point it implied their basic assumptions (77, p. 15).

This aptly catches James's influence on clinical psychology, not through work directly in the field or with the method, but through the fertile (and contradictory) character of his thinking.

In addition to the stimulation of his writings, James did take specific action of direct relevance to clinical psychology in his support of Clifford W. Beers, whose book, *A Mind that Found Itself* (19), did so much to further the mental hygiene movement. This he did through an endorsing letter which appeared in the first edition and, according to Henry James, his son, by departing from his fixed policy of "keeping out of Committees and Societies" (64, p. 273). In addition, he was interested in psychical research and in the efforts of Freud and Jung, although dubious about both of these trends (64).

Obviously it is impossible to capture the full flavor of William James in a paragraph or two, but this "defender of unregimented ideas" is at least the eccentric brilliant uncle of

the men in clinical psychology who followed after.

Another of the pioneers of this time and place was G. Stanley Hall. He was more influenced by the evolutionary concept stemming from Darwin than by French psychopathological thinking. Shakow, in considering Hall's influence on psychiatry, so well summarizes his contributions that they may be seen as contributions to clinical psychology as well. He writes that it was:

Hall, the propagandist, who gave Freud his first academic hearing, who gave courses in Freudian psychology beginning in 1908 and whose pressure for its consideration remained life-long; Hall, who influenced Cowles in establishing the psychological laboratory at McLean Hospital which had as directors following Hoch, Franz, Wells and Lundholm; Hall, who stimulated Adolf Meyer, by his early interest in child study, to write his first paper on a psychiatric topic—*Mental Abnormalities in Children during Primary Education . . .*—and who did so much to make the country child-conscious; Hall, whose students Goddard and Huey (also Meyer's students at the Worcester State Hospital) did the early pioneer work on feeble-mindedness . . . Hall, whose bravery in handling the problem of sex did so much to break down the first barriers, thus greatly facilitating the later child guidance handling of this and related problems; Hall, whose student Terman achieved so much in the development of the Binet method in the United States and whose student Gesell did so much for other aspects of developmental psychology; Hall, whose journals regularly published material of psychopathological interest; Hall, the ramifications of whose psychological influence are most pervasive in fields related to psychopathology . . . (92, p. 430).

Certain other factors might also be mentioned. Before his period as president of Clark University, Hall, while at Johns Hopkins, held weekly

clinics at Bay View Hospital and, until its medical staff was organized under his direction, served as lay superintendent. For a period of years he taught and demonstrated for psychiatrists at Worcester State Hospital, handing over the actual instruction in 1895 to Adolf Meyer, but continuing his interest in the field (74). Other students of this period who made substantial contributions to clinical psychology included Blanchard, Conklin, Kuhlmann, and Manteer.

Something of the spirit and activity of the associates of these men may be captured by an examination of the journal that was begun early in the century. *The Journal of Abnormal Psychology*, later called *The Journal of Abnormal and Social Psychology*, was a major source of publication of the more enlightened efforts of its time. Until 1913, when the *Psychoanalytic Review* was founded, it was the only journal in which psychoanalytic papers were published (32). Founded in 1906 for the express purpose of serving both medicine and psychology, it had as its editor Morton Prince, later professor of psychology at Harvard University, and numbered among its associate editors Hugo Münsterberg, James Putnam, August Hoch, Boris Sidis, Charles L. Dana, and Adolf Meyer. The papers in the first issue aptly catch the various influences at work in the psychology and psychiatry of the day. The first is a paper by Janet and thus represents the French psychopathological school; the second concerns hypnosis; the third, a critique of Freud by Putnam (the first article in English calling attention to Freud's work); and the fourth, a paper by Morton Prince concerning his most famous case of multiple personality, Miss Beauchamps. The

first book review in this new journal was that of Freud's *Psychopathology of Everyday Life*, which had been published in Germany in 1904. So far as this writer is aware, the first critical article concerning psychoanalysis by an American psychologist appeared in the February 1909 issue of this journal. It was entitled "An Interpretation of the Psychoanalytic Method in Psychotherapy with a Report of a Case so Treated" (90). This is apparently the second instance of a report of personal psychotherapeutic experience by a psychologist.<sup>2</sup> Its author, known for endeavors in fields far removed from this, was Walter Dill Scott, the psychologist, later president of Northwestern University.

The situation in the official psychiatric journal may be used for contrast. The first psychoanalytic paper to appear in the *American Journal of Insanity* was in the October 1909 issue. This paper was by Ernest Jones of Toronto and deplored the fact that Freud's methods had been neglected. None of Freud's books was reviewed in this journal for some years and, indeed, the first review to appear was that of Brill's *Psychoanalysis* in July 1914.

Isador Coriat (32), in presenting some reminiscences of psychoanalysis in Boston, attributes the interest in psychotherapy there to the stimulation of William James. Although A. A. Brill began psychoanalytic practice in New York in 1908, he was the only psychiatrist in the United States at that time engaging in such practice. He, Putnam, and Ernest Jones,

<sup>2</sup> Before taking a medical degree, Boris Sidis, then a psychologist, published in 1907 in the *Boston Medical and Surgical Journal* (96) a series of cases of what he called hypnoidal states treated by his particular method of suggestion.

then of Toronto, were the first in America to do active work with psychoanalytic methods. The first English translation of a work by Freud, *Selected Papers on Hysteria*, appeared in 1909 according to Coriat (32). It was in this same year that G. Stanley Hall, as president of Clark University, invited both Freud and Jung to come to the United States to lecture on the occasion of the twentieth anniversary of Clark University. Both by attendance and by the subsequent publication of these lectures in the *American Journal of Psychology* (51) psychologists became more familiar with their work. In the meantime, Brill (23) was translating Freud's works, and other psychoanalysts began practice. By 1911 there was enough interest that the first psychoanalytic association, the New York Psychoanalytic Society, was founded.

In view of these factors in the history of clinical psychology, it is possible to offer the interpretation that actually it was partly the psychologists and not psychiatrists alone, as is commonly supposed, who offered the first support to psychoanalysis in the United States. To be sure, in the twenties the psychiatrists in increasing numbers became interested and during the following twenty years became so firmly identified with the field that it is only today that psychologists, as psychologists, are again beginning to assume any prominence in psychoanalytic thinking and practice.

#### *The Psychologist and the Psychological Clinic*

It has been accepted by psychologists quite generally that the case leading to the founding of the first psychological clinic was treated by Lightner Witmer (114) at the University of Pennsylvania in March 1896.

Witmer was the first to speak of the "psychological clinic," of "clinical psychology," and the "clinical method in psychology" (26). The history of his clinic has been discussed elsewhere (26, 27, 93, 107, 114) and is quite well known. It is, therefore, unnecessary to dwell upon it. Instead, after very briefly examining its functioning, attention will be given to the extent of its influence upon the history of clinical psychology.

Even a cursory examination of the early issues of the *Psychological Clinic*, a journal founded and edited by Witmer, will show that the work attempted in this clinic included referral to medical sources, the presence of social workers, and many other "modern" innovations discussed by the writer elsewhere (107). On the other hand, although the juvenile court and social agencies referred cases to Witmer's clinic, the great majority came from the school system. Much attention was paid to the relation of physical defects and neurological conditions to behavior problems. Cooperation with special teachers of the blind and deaf and the mentally defective was stressed. In general, intellectual aspects of children's problems were emphasized, using a biographical approach. Relatively few psychologists published in the *Psychological Clinic* in the early years. Educators, either teachers, principals, or professors, wrote the majority of the articles during this period. In later years the publications of psychologists predominated. The articles are chiefly of antiquarian interest today.

The clinic founded by Seashore (91) at the University of Iowa about 1910 was modeled after Witmer's clinic, and others, such as that founded by J. E. W. Wallin of the

University of Pittsburgh in 1912, undoubtedly owe part of their impetus to it, but many other psychological clinics and activities seemed to grow up independently and with little knowledge of the development of this first clinic (97). For example, Seashore (91) speaks of his as the "second" psychological clinic. And yet in 1914 Wallin (105) found about 20 psychological clinics to be in existence, of which some at least must have developed under a different tradition except in the rather unlikely event that the great majority were founded after 1910, but before 1914. Although the Witmer clinic has been functioning continuously since its inception, it is quite difficult to find evidence of its effects upon clinical psychology today. This has not been due to lack of local support; rather it is because its influence did not spread beyond Philadelphia to any considerable degree. The reasons for this relative lack of influence will be discussed after considering a related development: the child guidance movement.

#### *The Psychologist in Child Guidance*

Still another stream which merged into the torrent that is clinical psychology today came from the so-called child guidance movement. In this effort William Healy (59), a psychiatrist, was the most important early figure. The beginnings of this movement arose from the conviction that antisocial behavior was treatable by psychiatric means. A subsequently discarded tenet which went hand in hand with this conviction was an emphasis upon pathology. Hence the first "child guidance" clinic, at the time of its founding in Chicago in 1909, was called "The Juvenile Psychopathic Institute." It is perhaps prophetic that the selec-

tion of Healy for the position of director was "as a pupil of James and a free lance in competition with a more rigid Wundtian and experimentally and statistically minded psychologist" (76, p. 242). Its first staff was very small, consisting of Dr. Healy, as psychiatrist, Dr. Grace M. Fernald, as psychologist, and one secretary. It is important to note that no social worker was a paid member of the staff, but Healy indicates that social workers from cooperating agencies worked with them from the very beginning. Only later did the specialty of psychiatric social worker, as such, emerge. Mental testing by Fernald, and later by Augusta F. Bronner, emphasized performance testing and other instruments of local origin. In 1910, however, Healy introduced the Binet-Simon tests into the United States (as did Goddard at Vineland simultaneously and independently). A direct outgrowth of the use of this and other instruments was the publication in 1927 of a *Manual of Individual Tests and Testing* (25) by Bronner, Healy, and their co-workers. Both Healy and Bronner had migrated eastward, organizing in 1917 a clinic in Boston under the name of the Judge Baker Foundation, later changed to the Judge Baker Guidance Center. This venture was enormously successful and resulted in still further important work in the field of delinquency. Many publications, including several books upon problems of the delinquent, had considerable influence upon patterns in this field.

In contrasting the relative success of Healy's venture and its continuity with the present with the relative lack of influence of Witmer's clinic, Shakow (92) presents a thoughtfully detailed statement, one or two points of which can be mentioned. The psy-

chologist Witmer was concerned with intellectual aspects of the functioning individual, worked primarily with mental defectives or school retardation problems, when concerned with medical aspects focused more on the physical or neurological, and, most important of all, identified himself with the Wundt-Kraepelin point of view. On the other hand, the psychiatrist Healy was concerned with affective aspects of the personality, worked primarily with behavior problems and delinquency, when concerned with medical problems stressed the psychiatric, and, again most important of all, was profoundly influenced by James and Freud. Although a pioneer, Witmer turned his back on almost all that was to predominate in the later days of clinical psychology and became of historical significance only. Healy is still a contemporary.

#### *The Psychologist in Mental Hospitals*

The importance of McLean Hospital in Waverly, Massachusetts has never been fully appreciated in the history of psychiatry and psychology. Founded in 1818, its superintendent at the turn of the century was Dr. Edward Cowles, a former surgeon in the Union Army.<sup>3</sup> Years later he took some incidental training in psychology at Johns Hopkins (57). In many ways he was a man ahead of his time. He encouraged research and brought to this hospital biochemists, pathologists, physiologists, and psychologists. One could date the beginnings of conjoint medicine as taking place at McLean Hospital since these approaches were used in its laboratory sometime before 1894. In that year Hall described the laboratory as fol-

<sup>3</sup> SHAFFER, P. A. Personal communication, 1952.

lows: "The work of this laboratory was begun in 1889, for the clinical purposes of the hospital. It is sought to combine neurological studies in the departments of psychiatry and physiological psychology, and their relations with anatomical and chemical pathology, etc." (57, p. 358). Only a quotation from Cowles will bring out the contemporary ring of his words:

The purpose of establishing and developing the laboratory has been carried on under much difficulty, naturally due to the newness of the attempt to combine with psychiatry the other departments of scientific medical research. The pathology of the terminal stages of insanity must be studied as heretofore, and it is necessary to add that of the initial conditions which lead to mental disorder. Such studies must therefore be combined with physiological psychology in the attempt to determine the exact nature and causes of departures from normal mental function. Also, in the dependence of these changes upon general physiological processes, and in order to take into account all the elements of vital activity involved, it is supremely necessary to study both physiological and pathological chemistry in their direct and indirect relations to mental changes (57, p. 363).

Research efforts along these lines apparently first emerged from this laboratory. In presenting the history of psychiatric research, Whitehorn (111) recognized this contribution and first described McLean Hospital and its work before dealing with any other developments.

Cowles, in a review of the progress in psychiatry at the time of the fiftieth anniversary of the American Psychiatric Association in 1894, emphasized the importance of what he referred to as the systems of "new psychology" as one of the "most hopeful signs of progress" to bring about advancement in the under-

standing of mental diseases (34). Either as frequent visitors from nearby Boston or as members of the staff of McLean Hospital at this time were Morton Prince, August Hoch, Boris Sidis, and Adolf Meyer. Interest in psychology is shown by the fact that Cowles and William Noyes, of the same hospital, were among the approximately 13 to 18 individuals who were present at the founding of the American Psychological Association at Clark University in 1892 (36).

In 1893 August Hoch (75) was selected by Cowles to be psychologist and pathologist at McLean. The use of the term *psychologist* was neither idle nor esoteric. Having previously received a medical education, he now was sent abroad for further training, and it would appear that much of his training was in psychology with Wundt, Kulpe, Marbe, and Kiesow. He also worked with Kraepelin. On assuming his post at McLean he turned to work with the ergograph in clinical problems and in the first volume of the *Psychological Bulletin* (62) summed up experimentation in this field. Subsequently, as professor of psychiatry at Cornell and director of the Psychiatric Institute, he turned to more narrowly psychiatric problems, but there would appear to be little doubt that during this period at McLean he functioned, in part at least, as a psychologist.

It was in this atmosphere that a psychological laboratory was founded. This laboratory was begun in 1904 at McLean Hospital by Shepard Ivory Franz (50). It was influential in the *rapprochement* of psychology and psychopathology, although often interested in matters more physiological than psychological. The laboratory became estab-

lished under the direction of Franz, and on his leaving for what is now St. Elizabeths Hospital of Washington, F. Lyman Wells was appointed his successor and remained there until 1921.

Franz continued his interest while in Washington, not only writing such articles with a modern ring, although published in 1912, as "The Present Status of Psychology in Medical Education and Practice" (25), but also introducing in 1907 a routine clinical psychological examination of all new patients in a mental hospital setting. This was probably the first instance of routine psychological testing of psychiatric hospital patients. Among Franz's associates during the early period were Grace H. Kent and Edwin G. Boring, both of whom published on learning in dementia praecox. Although Boring, as is well known, returned to other fields, he nevertheless felt that the summer he spent in the hospital was a very valuable, broadening experience (22). From 1906 to 1921 Grace H. Kent was psychologist at Philadelphia Hospital, Kings' Park State Hospital, and St. Elizabeths, respectively. In 1922 she went to Worcester State Hospital, remaining there until 1926 (79). Thereafter, for many years she was at Danvers State Hospital.

#### *The Psychologist and Institutions for the Mentally Defective*

It was Goddard's laboratory at the Vineland Training School that was the second center to be devoted to the psychological study of the feeble-minded.<sup>4</sup> Henry H. Goddard became director of psychological research at this institution in 1906 and was influential in the establishment of the

<sup>4</sup> In 1898 Wylie, a physician, began psychological testing at the state institution for the feeble-minded at Faribault, Minnesota (98).

psychologist as a person working with the mentally defective. As mentioned earlier, he first translated and used the Binet in this country. For practical purposes, the use of the Binet was at this time almost exclusively restricted to the feeble-minded. It was from this center that the Binet spread to other institutions (84). His directorship continued until almost the twenties.

#### *Psychology as a Profession*

It was as early as 1904 that Cattell (30) made the prediction that there would eventually be a profession as well as a science of psychology. Actually professional action preceded this pronouncement.

For purposes of this presentation the relevant characteristics of a profession include establishment of commonly agreed-upon practices concerning relationship with colleagues and with the public served. The questions of competency and the means of controlling competency immediately arise. Traditionally, a profession controls competency among its own members. Thus, self-determined control of its members is the hallmark of a profession.

The first stirrings of attempts at control arose in the American Psychological Association and took the form of considering control of clinical procedure through evaluation of test data. In 1895, only three years after the founding of the Association, J. Mark Baldwin, in the words of Fernberger, "proposed the formation of a committee to consider the feasibility of cooperation among the psychological laboratories for the collection of mental and physical statistics" (43, p. 42). The committee that was appointed, chaired by Cattell, called itself "The Committee on Physical and Mental Tests," but the battery

of tests they proposed for try-out to develop norms gained little acceptance so that after 1899 no further word was heard from this committee. Another committee for the purpose of establishing methods of testing was appointed in 1907 and continued until 1919. It made some progress, for example, sponsoring research on the Woodworth-Wells Association Tests, but it fell far short of the ostensible goal.

In 1915, on the motion of Guy M. Whipple, the Association went on record as "discouraging" the use of mental tests by unqualified individuals. In 1917 a committee to consider qualifications for psychological examiners was appointed, and two years later one to consider certifying "consulting" psychologists. In 1919 the Section of Clinical Psychology within the American Psychological Association was formed (43). In large measure, it was a "special interest group" concerned with arranging programs at the annual meetings and the like. Its members were, however, drawn into the discussion, pro and con, of the merits of certification. After much maneuvering, favorable action on certification of clinical psychologists finally resulted, and the first certificates were granted after the 1921 meeting. However, only twenty-five psychologists applied, and the project was abandoned. The death blow was dealt by an APA policy committee which considered that certification was not practicable and, on vote of the APA membership in 1927, discontinued certification. In some measure at least, the decision was influenced by the realization that with certification went the problem of enforcement of the standards instituted, especially on psychological workers outside the membership. Thereafter, according to Fernberger (43), there was a period of some years

without important action within the American Psychological Association on these problems.

Internship training, as distinguished from academic course work, is a manifestation of professional training. Morrow (78) indicates that Lightner Witmer was apparently the first to suggest practical work for the psychologist through training school and laboratory. However, the first actual internships were those offered by the Training School in Vineland, New Jersey, under the supervision of H. H. Goddard. This program began in 1908 and has continued down to the present time. In 1909 William Healy began accepting graduate students at the Juvenile Psychopathic Institute in Chicago, while the first internship in a psychiatric institution for adults was established in 1913 at the Boston Psychopathic Hospital under the direction of Robert M. Yerkes. Other earlier internships include those at Worcester State Hospital, McLean Hospital, the Western State Penitentiary in Pennsylvania, and the New York Institute for Child Guidance.

#### CLINICAL PSYCHOLOGY IN THE TWENTIES AND THIRTIES

In the twenties and thirties clinical psychology left the period of its lusty, disorganized infancy and entered its rather undernourished but rapid and stormy adolescence. As late as 1918 only 15 members or 4 per cent of the APA listed the field of clinical psychology as a research interest (44). This rose to 99 members or 19 per cent in 1937. In that year the newly instituted membership category of Associate showed 428 or 28 per cent interested in clinical psychology, the largest field of interest for this class of membership. In increasing numbers clinical psychologists were employed in hospitals,

clinics, schools, penal institutions, social agencies, homes for the feeble-minded, industrial plants, and the entire gamut of agencies concerned with human welfare. For example, Finch and Odoroff (46), in a survey concerning employment trends, indicate that of 1,267 members of the American Psychological Association in 1931, 286 or 26.9 per cent were not in teaching positions. In 1940 the number of nonteachers had swelled to 888 or 39.3 per cent of the membership. Clinical nonteachers increased from 95 to 272 during this ten-year period.

It was during this period that many psychologists did yeoman service for clinical psychology without being primarily identified with the field. Carl E. Seashore (91) may be used as an illustration. It has already been noted that he founded a psychological clinic at the University of Iowa about 1910. During the period now under consideration he was interested in the relationship between psychology and psychiatry and took the lead in organizing a national conference on this topic. He also aided in founding the Iowa Psychopathic Hospital and worked with Samuel Orton, Edward Lee Travis, and Wendell Johnson in speech pathology. Many other men such as Gardner Murphy, Goodwin Watson, Horace B. English, Albert T. Poffenberger, Kurt Lewin, Carney Landis, Robert M. Yerkes, Walter R. Miles, Gordon W. Allport, and Kurt Goldstein, although primarily associated with some other aspect of psychology, also performed services for the clinical field.

In spite of such developments as those just described, Loutitt (70) could indicate during the same period that "American Psychology, generally speaking, has not been greatly interested in practical problems of

human behavior" (70, p. 361). This contention applied to clinical psychology with as much force as, or more than, it did to other applications of psychology. Most of the difficulties that clinical psychology went through during this period as it groped toward professional stature were internal to the field itself. Both rapid growth and some hostility from the dominant entrenched forces in psychology are imbedded in the history of the period and influence many of the specific developments now to be discussed.

#### *The Psychometric Tradition*

The period of the twenties was, in the words of Merrill, a "plateau . . . [following] the initial impetus given to testing when these first tools of the clinician were being subjected to evaluation, and the exaggerated expectations of over-enthusiastic users were being reduced in the crucibles of research" (73, p. 283). Studies of validity, investigations of the constancy of IQ, application of the tests to new populations, studies of individual differences, the nature-nurture controversy, racial differences, the development of group testing, achievement tests, interest measures, and personality testing of the questionnaire variety occupied this and the following decade and helped to consolidate the gains of the previous period. Theories of intelligence and factor analysis are also intimately related to this trend. It was a period, as the term plateau implies, of masked gain which prepared the way for the present period.

More and more objections began to be raised to the limitations entailed by this approach. The development of group tests during and after World War I placed a premium on easy reproduction, rigid standardization down to the slightest detail,

and emphasis on the score obtained to the exclusion of all else. Measures of personality with these same characteristics were developed during the twenties and thirties. To some psychologists the results obtained were considered disappointing and sterile.

In 1927 F. L. Wells published *Mental Tests in Clinical Practice* (108), in which he stated vividly the major objection to a rigid psychometric approach:

An intelligent South Sea Islander, observing a psychometric examination, would be likely to regard it as a magic rite designed to propitiate friendly spirits in the patient's behalf. Should he observe a conscientious examiner in the apprentice stage, tightly clinging to forms prescribed, his idea would be confirmed, for none knows better than himself how slight a departure from the required formulae will not only destroy their beneficence but may well deliver the hapless sufferer into the hands of the malignant ghosts. Over against such esoteric views of psychometric methods is the customary and pragmatic one. The function of psychometrics is not the accomplishment of a ritual, but the understanding of the patient. The ceremony of mental tests is valuable so far as it serves to reach this end. When it fails, or stands in the way of doing this, proper technique demands that it be modified. Ability to do this intelligently is what distinguishes the psychologist, properly so called, from the "mental tester" (108, p. 27).

Further objections to exclusive reliance upon a psychometric approach arose with the emergence of projective techniques as an aspect of the dynamic tradition next to be considered.

#### *The Dynamic Tradition*

Many of the present developments in clinical psychology—the emphasis on understanding of personality func-

tioning, the attempt to relate present behavior to experiences of which the patient is unaware, the evaluative use of incidental verbalizations and physical behavior of the patient, and the artistic element in psychodiagnostic appraisal—stem in large measure from the dynamic tradition.

In terms of the sources of these influences, Sigmund Freud, of course, looms largest. He and his fellow analysts profoundly affected the thinking of many clinical psychologists, who were for the most part passive recipients of this influence. No longer did they share leadership with their medical colleagues as during the first twenty years of the century. The influence of psychoanalysis was felt directly on three of the specific manifestations of the dynamic tradition directly involving the psychologist—projective techniques, the Harvard Psychological Clinic, and the American Orthopsychiatric Association. The first, an approach to personality, the second, a clinic, and the third, a professional organization, share responsibility as the most important manifestations of the dynamic tradition in psychology of the day. Each will be considered in turn.

Hermann Rorschach, a Swiss psychiatrist, published with Oberholzer on the specific but intricate relationships which exist between his inkblot technique and psychoanalysis. The technique itself occupied much of his time between 1911 and his untimely death in 1922. In the United States pioneering work with the Rorschach was done by David M. Levy, a child psychiatrist, with whom Samuel J. Beck became associated beginning in 1927. In 1930 Beck presented the first Rorschach study in this country as his doctoral dissertation, and during the thirties the Rorschach technique came more and more into

prominence in clinical circles. Along with Beck, pioneer American psychologists who made major contributions to Rorschach literature during this period were Bruno Klopfer, Marguerite Hertz, and Zygmunt Piątrowski.

Two reasons are given by Beck (18) for the increasing preoccupation of psychologists with the Rorschach technique as compared to psychiatrists. There is, first, the division of labor with the Rorschach as one of the diagnostic testing instruments and, second, the fact that its use spread outside the narrowly psychiatric area into the schools, work with delinquents, industry, and the like. Other projective techniques, notably the Thematic Apperception Test, also appeared during the thirties.

In the meantime the psychodynamic emphasis began to be a part of the intellectual armamentarium of the psychologist. The article by L. K. Frank, "Projective Methods for the Study of Personality" (48), published in 1939, offered a rationale for the projective approach and stimulated both research and theoretical efforts in the decades to follow. Merrill summarizes other reasons for the rapid spread of projective testing as having

... had to do with significant changes that have been occurring in the clinician's self-concept and his changing perception of his role as his social responsibilities grow and expand. Projective tests have become important tools for the psychotherapists. These tests command the attention and respect of our colleagues in the medical fraternity, the psychiatrists. They constitute moreover, the basic technological structure upon which is being built a new systematic point of view, projective psychology with its own theory of personality. This new projective psychology has been aptly characterized as a psychology of protest. As both be-

haviorism and Gestalt psychology came about as protests against the established psychologies called structural, so this emerging projective psychology runs sharply counter to the traditions that have characterized individual psychology in America. Having something to push against, it can move (73, p. 286).

In 1927 the Harvard Psychological Clinic was founded by Morton Prince. Its express purpose was to bring together academic and clinical psychology. Henry A. Murray took over headship of the clinic early in the thirties and with a large group of collaborators, including Donald W. MacKinnon, Saul Rosenzweig, R. Nevitt Sanford, and Robert W. White, carried on a brilliant research project in personality functioning. This culminated in 1938 in the well-known *Explorations in Personality* (81).

The American Orthopsychiatric Association is an organization with ties to child guidance in particular and to the dynamic tradition in general. It was founded in 1924 with many of the leaders of the child guidance movement present (71). William Healy was elected president during this year and served through 1926. Later presidents included Karl Menninger, David Levy, and in 1931, the first psychologist to be president, Augusta F. Bronner. Other psychologist presidents were Edgar A. Doll, Samuel Beck, and Morris Krugman. After thinking through the problem of membership, originally restricted to psychiatrists, the pattern emerged in 1926 of having as full members "psychiatrists, psychologists, and other professional persons whose work and interests lie in the study and treatment of conduct disorders" (71, p. 199). Both the letter and spirit of this method of organization for work interchange, support, and

advance have continued to the present day. However, there has never been any question, as might have been foretold from the original organization, but that psychiatrists were dominant in it. For example, twenty-one of the first twenty-six presidents held the M.D. degree, only four being psychologists and only one a social worker. This organization continues to wield much influence both through its journal, *The American Journal of Orthopsychiatry*, and through its annual meetings which are characteristically attended by far more nonmembers than members.

#### *Psychological Clinics*

This was a period during which psychological clinics reflected the plateau of the psychometric tradition. Some new clinics appeared; others closed their doors (70). In 1934 a survey report of a questionnaire of psychoeducational clinics by Witty and Theman (115) appeared. On the basis of their returns they estimated that there were about 50 such clinics. This figure may be contrasted with the approximately 20 found by Wallin in 1914 (105). In 1932 the median length of time the clinics had been in existence was four years. Located in colleges, universities, teachers colleges, and normal schools, their stated purposes involved (a) providing schools, social agencies, and individuals with diagnostic test services and remedial methods in order to bring about educational, vocational, and social adjustment; (b) training students in giving and interpreting tests; and (c) research with emphasis on the study of deviates, causes and treatment of learning difficulties, and work with remedial materials. It would appear that this survey epitomizes the work of psychological clinics of the day, featuring

emphasis on testing and remedial education.

#### *Child Guidance*

The period 1922-1927 was one in which the National Committee for Mental Hygiene on behalf of the Commonwealth Fund established demonstration clinics in a variety of cities and rural areas for the purpose of showing both their need and the work they could do (99). For the first time they were called "child guidance clinics." Eight clinics were permanently established directly as a result, and many others were at least partially stimulated by this effort. It was the announced intention from the very beginning that eventually expenses for their maintenance would be absorbed by the community in which they were located. Deliberate experimentation as to method of organization in relation to other agencies was carried out—some were attached to the courts, others to local charities, to university and to teaching hospitals. The child guidance clinic plan of organization called for the professional personnel to include at least a psychiatrist, a psychologist, and a social worker. These activities in their formative stages continued roughly over the decade 1920-1930. To be sure, their influence and organization continued thereafter, but this period marked the heyday of their unique contribution.

An important shift of focus of attention had been occurring during this period. No longer was the delinquent of primary interest. Nor was there much concern with mental defectives, epileptics, or neurological cases. Instead, emphasis was placed upon maladjustment in school and home, especially that centering around parent-child relationships.

The clinics began to concentrate upon problems of the individual who may be spoken of as falling within the normal range of intelligence, the roots of which may in some measure be traced to emotional difficulties.

#### *The Psychologist in Mental Hospitals*

In 1921 Wells left McLean Hospital for Boston Psychopathic Hospital where he served as head psychologist until 1938. This hospital also became a center of clinical activity and training. A pioneer in present-day clinical psychology, David Shakow, now of the Illinois Neuropsychiatric Institute, is still very active. For a period of nearly twenty years, Shakow served as director of psychological research at Worcester State Hospital. His activities, along with his research efforts, included direction of an internship training program. It apparently was the closest in spirit to the modern internship program, and his experience derived in this setting was of great value in formulating present-day practices concerning internship.

A gradual increase in the number of clinical psychologists in mental hospitals was taking place. However, the geographical isolation of most such hospital psychologists apparently accentuated an isolation on other grounds so that the effect of this aspect of the development of clinical psychology was not as important as it was to be in the decades to come. Nevertheless, some psychologists were beginning to suspect that their approach was unduly limited. As a result, there were serious attempts at broadening the scope of testing efforts, to escape the atomistic tradition by means of research and theorizing concerning the personality of their patients.

#### *Mental Deficiency*

In 1919 Goddard was succeeded by Stanley D. Porteus as director of the Vineland Laboratory. Under both his direction and the subsequent direction from 1925 until 1949 of Edgar A. Doll the clinical problems of feeble-mindedness received intensive and extensive study.

It was precisely in the field of intelligence testing that clinical psychology was most advanced during this period. Psychometric testing of suspected mental deficiency was widely accepted, and the psychologist was the authority in this field (28). Nevertheless, as Buck (28) indicates, appreciation of the complexity, rather than the simplicity, of the diagnosis of mental deficiency emerged during these two decades.

The fact that a person was doing clinical psychological work with mental defectives unfortunately indicated almost nothing about the nature of the training and experience of the practitioner in question. In 1940 Hackbusch (56) reported an inquiry concerning psychological work in state and private institutions for the mentally defective. Of the approximately 100 institutions which replied apparently all were doing some form of psychological testing. However, less than half had a psychologist on their staff. The remainder had their testing done by outside sources, or by teachers, social workers, and physicians on their own staffs. It is also noteworthy that the "psychologists" apparently varied widely in the nature of their background. Some had less than an A.B. degree, while others had an A.B. or an M.A. in addition, but very few held the Ph.D. degree. Therefore, despite the acceptance of their work, the status of psychologists and psychological work in the twen-

ties and thirties was somewhat confused.

#### *Psychology as a Profession*

The origins of professional activity, as has been indicated, were centered within the American Psychological Association. This period extended from 1895 to the mid-twenties. Founded to advance psychology as a *science*, the Association had not been singularly successful in reflecting the interests of its members either in applications of psychology or in their professional aspirations. The twenties and thirties were characterized by the advent of other organizations more directly concerned with professional problems.

In 1917 a group of psychologists interested in the advancement of the practice of psychology met in Pittsburgh, Pennsylvania. Leta S. Hollingworth took the initiative in bringing the group together, and prominent charter members included Bronner, Fernald, Healy, Kuhlmann, Pintner, Terman, Whipple, Wells, and Yerkes. To quote Symonds, "After a brief history of two years, during which a bitter struggle went on in the American Psychological Association over the question of authority for certification of psychologists for clinical work, the American Association of Clinical Psychologists became defunct through the adoption by the APA of a report recommending the establishment of the AACP as a Section of Clinical Psychology" (100, p. 337). According to the same writer, the next step was the slow development of various local groups concerned with applied and professional matters in several states.

In 1930 the Association of Consulting Psychologists was reorganized from a still earlier association founded

in 1921 (40, 52). Gradually it extended its membership beyond New York and environs and became one of the more important elements later to merge into the American Association for Applied Psychology (AAAP). The organization meeting of this association took place in 1937. Many of the difficulties in organizing centered upon the standards for membership. Then, as now, there was the dilemma of maintaining standards and yet not setting them so high as to exclude the majority of those doing work in the applied fields. Eventually this was settled, and a national organization concerned with all aspects of the application of psychology came into being and became the dominant national professional organization. A divisional structure was followed with clinical, educational, industrial, and consulting sections.

*The Journal of Consulting Psychology* was at first a publication of the Association of Consulting Psychologists and then of the AAAP. Papers in clinical, educational, industrial, and consulting psychology appeared, but a considerable portion of space was devoted to organizational and professional matters (100).

Thus, there existed at the close of the thirties two major psychological societies—one dedicated to the advancement of psychology as a science and the other to its application. Generally speaking, members of the latter also had membership in the former but sincerely felt the essential nature of their applied organization. The Psychometric Society and the Society for the Psychological Study of Social Issues were also founded during this period. In part at least, these organizations arose because of similar dissatisfaction with the adequacy of representation of some of

their interests in the American Psychological Association. So the thirties closed with at least the possibility of dangerous rifts in the ranks of psychologists. However, as is well known, this danger passed in the forties with all of these organizations integrated into the reorganized American Psychological Association (116). In 1945 this reorganization went into effect. Both in spirit and in practice the American Psychological Association represents psychology as a science and as a profession.

#### *The Psychologist and Therapy*

During the twenties and thirties there appeared to be a gradual increase in the number of clinical psychologists engaging in therapy. From the time of Sidis and Scott at the turn of the century some psychologists had been so employed. In many instances psychotherapeutic practice grew out of the psychologist's educational function. Considered as expert both in matters of learning as a subject of investigation and in education as a field of endeavor, the psychologist worked with patients, particularly children, with whom remedial education was necessary. A similar process took place to a lesser extent in psychiatric clinics. It was in the hospitals that this development lagged, partly because the sheer press of numbers of patients confined the psychologists to psychodiagnostic tasks and partly because psychotherapy, except at a few institutions, was not practiced at all.

There was relatively little difficulty in interprofessional relations with psychiatry during this period. In large measure this was because there were few psychologists practicing therapy, and these few were doing so under institutional auspices and exceptional circumstances. Then too,

the psychiatrist himself was more isolated both from his medical colleagues and from the public than he is today. More concerned with the psychotic and the adult than with the neurotic and the child, his path did not as often cross that of the psychologist as it did in the forties and fifties.

No continuity in the development of psychotherapists among psychologists is discernible from generation to generation. Neither Sidis nor Scott stimulated psychologists to work with psychotherapeutic problems. In later years individual psychologists prominent in psychotherapy gained in stature, not unaided to be sure, but also not from the combined efforts of any group or from the work of one senior individual. Phyllis Blanchard, for example, an acknowledged leading therapist, as attested to by her presence in leading symposia and by books on the topic, neither received her training in therapy from psychologists nor participated in the training of psychologists in therapy. Other therapist-psychologists, also, developed along individual lines. The work of Carl Rogers, although begun in the thirties, did not reach national prominence until the forties.

#### PSYCHOLOGISTS IN THE ARMED SERVICES AND THEIR EFFECT UPON PSYCHOLOGY IN THE POSTWAR PERIOD<sup>6</sup>

About 1,500 psychologists served in the armed services during World War II. About one out of four psychologists thus was called upon to function in an applied field—that is, psychology applied to the very

<sup>6</sup> This section of the article is a modification of a section of a chapter in a book edited by the writer (106). The permission of Harper and Brothers, publishers, to include this section is acknowledged gratefully.

practical problem of war. Moreover, this group was predominantly young, averaging about 32 years of age (24), thus including many individuals just reaching professional maturity. It is not unduly optimistic to suppose that some of their experiences during these tours of duty carried over in attitude and practice to the postwar years.

To appreciate properly certain changes of attitude, it must be remembered that a considerable number of psychologists in uniform were products of an academic tradition whose isolationist tendencies in regard to professional application prior to the war they were quite willingly and even complacently furthering. In fact, Andrews and Dreese (16) found that almost 90 per cent of the psychologists in military service were in academic or governmental work prior to the war.

From the process of learning to apply their psychological training to the military situation, later consideration revealed at least two major trends that have had, and will continue to have, profound effect upon contemporary psychology. They discovered to their mild surprise, and to the considerable amazement of their colleagues from other disciplines, that their general training in psychological methods was capable of application to many problems which at first seemed utterly alien to their background. From aircraft instrument-panel design to selecting underwater demolition teams, psychologists found that they, in collaboration with specialists from other fields, had something valuable to contribute. Realization was forced upon them that an experimental background in psychology is capable of transfer to intelligent and capable handling of many sorts of problems.

Paradoxically, however, they gained added respect for the clinical approach. In this connection it must be realized that almost half of the psychologists used clinical and counseling procedures during some part of their period in uniform (24). Many psychologists, willy-nilly, were placed in a position where they functioned in selection and assignment, sat as members of discharge boards, worked as members of clinical teams, conducted therapeutic sessions, both group and individual, and in these and many other ways used diagnostic and treatment methods. Concrete expressions of this interest can be found in an article by Britt and Morgan (24) concerning the results obtained from a questionnaire mailed to every psychologist in uniform. They conclude that there was an overwhelming interest in having more practical postwar graduate training. Nearly 24 per cent of the suggestions for new courses for graduate study were clearly within the general clinical field. At least some of the armed services psychologists who had previously not been particularly receptive came to understand and appreciate the contributions, past and potential, of the clinical method. This impression is verified by the finding in a survey by Andrews and Dreese (16) that three times as many military psychologists engaged in clinical work after the war as had done so in the prewar period.

#### CLINICAL PSYCHOLOGY TODAY

With the coming of the forties and World War II, one leaves the realm of history and enters the present. It would be both hazardous and presumptuous to attempt to trace in detail the events from this time on. Nevertheless, certain factors in the foregoing account may be related to

present trends. Clinical psychology as a method, as an attitude, and as a field of endeavor is reflected in its past.

It would appear that clinical psychology and academic psychology have influenced each other markedly, with a reciprocal, symbiotic relationship having been formed. Other disciplines, notably the medical and particularly the psychiatric and psychoanalytic, influenced and vitalized clinical psychology.

Through the thirties certain predominant aspects may be referred to as "child," "psychological," and "clinical" as contrasted with "adult," "psychiatric," and "institutional" functions. Distinctively psychological clinics and work with children are not only important because of their service and scientific value, but also for the community orientation that they manifest and the preventive emphasis that they maintain. And yet since the thirties the emphasis has shifted.

The "adult," "psychiatric," and "institutional" aspects of clinical psychology appear to be dominant today, but this is by no means an unmixed blessing. Many of the more vocal leaders of the field, including to some extent the official committees of the American Psychological Association, have fostered emphasis upon the former. The extremely valuable support rendered by the Veterans Administration to our training and practice has emphasized the current trend. Work with adults in a psychiatrically oriented institution is a specialty, albeit an important one, in the broader field.

With the forties also came the domination in the history of clinical psychology of one of the trends previously sketched. This was the emergence and implementation of a

concept of a profession of psychology. One illustration will suffice. Until after World War II, there was relatively little demonstrable agreement about the training, nature, duties, or status of the clinical psychologist. To quote Eysenck, "A person who called himself a 'clinical psychologist' might be someone of great eminence, highly qualified academically and with 20 or 30 years of practical experience in the fields of diagnostic testing, research, and therapy, or he might be a student just graduated from the University, without any kind of relevant experience, capable only of grinding out Binet I.Q.'s without even an adequate understanding of their relevance to the clinical problem presented" (41, p. 711). The facet of the professionalization of a psychologist, although not completely defined today, has reached a degree of precise formulation undreamed of a few years ago.

Current issues and accomplishments, stabilizing trends, and unresolved problems may be related to the emergence of psychology as a profession. Factors making for the present stabilization include the agreement of the great majority of interested parties concerning diagnostic appraisal as a task of the clinical psychologist (7, 9, 12, 55, 94), the present organization and function of the American Psychological Association (4, 116), current efforts directed toward the training of clinical psychologists (9, 10, 12, 14), present activities looking toward codification of ethical problems (5, 6, 17, 21, 61), and the influence of such institutions as the American Board of Examiners in Professional Psychology (2, 3), state societies (11), the United States Public Health Service (42, 104), the Veterans Administration (1, 58, 103),

and the armed services (102). On the other hand, currently unresolved issues face the profession today. The problems on which there are differences of opinion both in psychology and in other professions include psychotherapy as a task of the psychologist and the nature of the relation of psychology to psychiatry and medicine (7, 13, 54, 55, 72, 87, 94), the nature of the relation of psychology to social work (33), the question of the advisability of certification and licensure (31, 47, 60, 109, 112, 117), the question of the desirability of private practice (38, 39), the position and function of non-Ph.D.'s in clin-

ical psychology (15, 20, 35, 67, 69, 98), and the "imbalance" in psychology between scientific and professional demands (63, 87, 88, 89). Not only do these problems have roots in the past, but they are also an expression of the period of professionalization of large segments of psychology today.

World War II focused the needs and demonstrated what could be done in clinical psychology; the period after the war is still feeling the pressure of these social needs and is witnessing the reactions, adaptive and otherwise, of a beginning profession to these demands.

#### REFERENCES

1. ADLER, M. H., FUTTERMAN, S., & WEBB, R. Activities of the mental hygiene clinics of the Veterans Administration. *J. Clin. Psychopath.*, 1948, 9, 517-527.
2. AMERICAN BOARD OF EXAMINERS IN PROFESSIONAL PSYCHOLOGY, INC. *Official Bulletin*. 1948, No. 1.
3. AMERICAN BOARD OF EXAMINERS IN PROFESSIONAL PSYCHOLOGY. The work of the American Board of Examiners in Professional Psychology: annual report of the Board to the members of the APA. *Amer. Psychologist*, 1951, 6, 620-625.
4. AMERICAN PSYCHOLOGICAL ASSOCIATION. By-laws for the American Psychological Association (as amended through September, 1951). In *Directory, American Psychological Association*. Washington, D. C.: American Psychological Association, 1951.
5. AMERICAN PSYCHOLOGICAL ASSOCIATION. *Ethical standards for psychologists*. Vol. 1: *The code of ethics*. Washington, D. C.: American Psychological Association, 1952.
6. AMERICAN PSYCHOLOGICAL ASSOCIATION. *Ethical standards for psychologists*. Vol. 2: *Source book of ethical problems, incidents, and principles*. Washington, D. C.: American Psychological Association, 1952.
7. AMERICAN PSYCHOLOGICAL ASSOCIATION, *Ad Hoc Committee on Relations Between Psychology and the Medical Profession*. Psychology and its relationships with other professions. *Amer. Psychologist*, 1952, 7, 145-152.
8. AMERICAN PSYCHOLOGICAL ASSOCIATION, COMMITTEE OF CLINICAL SECTION. I. The definition of clinical psychology and standards of training for clinical psychologists. II. Guide to psychological clinics in the United States. *Psychol. Clin.*, 1935, 23, 2-140.
9. AMERICAN PSYCHOLOGICAL ASSOCIATION, COMMITTEE ON TRAINING IN CLINICAL PSYCHOLOGY. Recommended graduate training program in clinical psychology. *Amer. Psychologist*, 1947, 2, 539-558.
10. AMERICAN PSYCHOLOGICAL ASSOCIATION, COMMITTEE ON TRAINING IN CLINICAL PSYCHOLOGY. Annual report of the Committee on Training in Clinical Psychology. *Amer. Psychologist*, 1951, 6, 612-617.
11. AMERICAN PSYCHOLOGICAL ASSOCIATION, CONFERENCE OF STATE PSYCHOLOGICAL ASSOCIATIONS. *CSPA Newsletter*, April, 1952. (Mimeo.)
12. AMERICAN PSYCHOLOGICAL ASSOCIATION, CONFERENCE ON GRADUATE EDUCATION IN CLINICAL PSYCHOLOGY, Boulder, Colorado. *Training in clinical psychology*. New York: Prentice Hall, 1950.
13. AMERICAN PSYCHOLOGICAL ASSOCIATION, DIVISION OF CLINICAL AND ABNORMAL PSYCHOLOGY, COMMITTEE ON PSYCHO-

THERAPY. Report. *Newsletter, Div. clin. abnorm. Psychol.*, 1950, 4, No. 2, Suppl. (Mimeo.)

14. AMERICAN PSYCHOLOGICAL ASSOCIATION, EDUCATION AND TRAINING BOARD. Doctoral training programs in clinical psychology. *Amer. Psychologist*, 1952, 7, 158.
15. AMERICAN PSYCHOLOGICAL ASSOCIATION, POLICY AND PLANNING BOARD. Annual report: 1951. *Amer. Psychologist*, 1951, 6, 531-540.
16. ANDREWS, T. G., & DREESE, M. Military utilization of psychologists during World War II. *Amer. Psychologist*, 1948, 3, 533-538.
17. Anon. Discussion on ethics: a little recent history. *Amer. Psychologist*, 1952, 7, 426-428.
18. BECK, S. J. Rorschach's test in this anniversary year. In L. G. Lowrey (Ed.), *Orthopsychiatry, 1923-1948: retrospect and prospect*. New York: American Orthopsychiatric Association, 1948. Pp. 422-455.
19. BEERS, C. W. *A mind that found itself*. New York: Longmans Green, 1908.
20. BLACK, J. D. A survey of employment in psychology and the place of personnel without the Ph.D. *Amer. Psychologist*, 1949, 4, 38-42.
21. BOBBITT, J. M. Some arguments for a code of ethics. *Amer. Psychologist*, 1952, 7, 428-429.
22. BORING, E. G. Edwin Garrigues Boring. In E. G. Boring, H. S. Langfeld, H. Werner, & R. M. Yerkes (Eds.), *A history of psychology in autobiography*. Vol. IV. Worcester: Clark University Press, 1952. Pp. 27-52.
23. BRILL, A. A. Introduction. In A. A. Brill (Ed.), *The basic writings of Sigmund Freud*. New York: Modern Library, 1938. Pp. 3-32.
24. BRITT, S. H., & MORGAN, JANE D. Military psychologists in World War II. *Amer. Psychologist*, 1946, 1, 423-437.
25. BRONNER, AUGUSTA F., HEALY, W., LOWE, GLADYS M., & SHIMBERG, MYRA E. *A manual of individual mental tests and testing*. Boston: Little Brown, 1927.
26. BROTEMARKLE, R. A. (Ed.) *Clinical psychology: studies in honor of Lightner Witmer*. Philadelphia: Univer. of Pennsylvania Press, 1931.
27. BROTEMARKLE, R. A. Clinical psychology 1896-1946. *J. consult. Psychol.*, 1947, 11, 1-4.
28. BUCK, J. N. The present and future status of the psychologist in the field of mental deficiency. *Amer. J. ment. Def.*, 1949-1950, 54, 225-229.
29. CATTELL, J. M. Mental tests and measurements. *Mind*, 1890, 15, 373-381.
30. CATTELL, J. M. Retrospect: psychology as a profession. *J. consult. Psychol.*, 1937, 1, 1-3; 1946, 10, 289-291.
31. COMBS, A. W. A report of the 1951 licensing effort in New York State. *Amer. Psychologist*, 1951, 6, 541-548.
32. CORIAT, I. H. Some personal reminiscences of psychoanalysis in Boston: an autobiographical note. *Psychoanal. Rev.*, 1945, 32, 1-8.
33. COWAN, E. A. Correspondence. *J. consult. Psychol.*, 1945, 9, 64-65.
34. COWLES, E. Progress during the half century. *Amer. J. Insanity*, 1894, 51, 10-22.
35. DARLEY, J. G., ELLIOTT, R. M., HATHAWAY, S. R., & PATERSON, D. Are psychologists without Ph.D. degrees to be barred from membership in the APA? *Amer. Psychologist*, 1948, 3, 51-53.
36. DENNIS, W., & BORING, E. G. The founding of the APA. *Amer. Psychologist*, 1952, 7, 95-97.
37. DOLL, E. A. (Ed.) *Twenty five years: a memorial volume in commemoration of the 25th anniversary of the Vineland Laboratory, 1906-1931*. (Publ. Ser. 1932.)
38. ELLIS, A. The psychologist in private practice and the good profession. *Amer. Psychologist*, 1952, 7, 129-131.
39. ELLIS, A. (Chm.) Report of the Committee on Private Practice. *Newsletter, Div. clin. abnorm. Psychol.*, 1952, 5, No. 4. (Mimeo.)
40. ENGLISH, H. B. Organization of the American Association of Applied Psychologists. *J. consult. Psychol.*, 1938, 2, 7-16.
41. EYSENCK, H. J. Function and training of the clinical psychologist. *J. ment. Sci.*, 1950, 96, 710-725.
42. FEDERAL SECURITY AGENCY. *National Mental Health Act. Five years of progress, 1946-1951*. Washington, D. C.: 1951. (Mimeo.)
43. FERNBERGER, S. W. The American Psychological Association: a historical summary, 1892-1930. *Psychol. Bull.*, 1932, 29, 1-89.
44. FERNBERGER, S. W. The scientific interests and scientific publications of the members of the American Psychologi-

cal Association. *Psychol. Bull.*, 1938, **35**, 261-281.

45. FERNBERGER, S. W. The American Psychological Association, 1892-1942. *Psychol. Rev.*, 1943, **50**, 33-60.

46. FINCH, F. H., & ODOROFF, M. E. Employment trends in applied psychology. *J. consult. Psychol.*, 1941, **5**, 275-278.

47. FOWERBAUGH, C. C. Legal status of psychologists in Ohio. *J. consult. Psychol.*, 1945, **9**, 196-200.

48. FRANK, L. K. Projective methods for the study of personality. *J. Psychol.*, 1939, **8**, 389-413.

49. FRANZ, S. I. The present status of psychology in medical education and practice. *J. Amer. med. Ass.*, 1912, **58**, 909-911.

50. FRANZ, S. I. Shepard Ivory Franz. In C. Murchison (Ed.), *A history of psychology in autobiography*. Vol. II. Worcester: Clark Univer. Press, 1932. Pp. 89-113.

51. FREUD, S. The origin and development of psychoanalysis. *Amer. J. Psychol.*, 1910, **21**, 181-218.

52. FRYER, D. (Chm.) The proposed American Association for Applied and Professional Psychologists. *J. consult. Psychol.*, 1937, **1**, 14-16.

53. GALTON, F. *Inquiries into human faculty and its development*. London: J. M. Dent, 1883.

54. GREGG, A. The profession of psychology as seen by a doctor of medicine. *Amer. Psychologist*, 1948, **9**, 397-401.

55. GROUP FOR THE ADVANCEMENT OF PSYCHIATRY, COMMITTEE ON CLINICAL PSYCHOLOGY. The relation of clinical psychology to psychiatry. *Amer. J. Orthopsychiat.*, 1950, **22**, 346-354.

56. HACKBUSCH, FLORENTINE. Responsibility of the American Association on Mental Deficiency for developing uniform psychological practices in schools for mental defectives. *Amer. J. ment. Def.*, 1940-41, **45**, 233-237.

57. HALL, G. S. Laboratory of the McLean Hospital. *Amer. J. Insanity*, 1894, **51**, 358-364.

58. HAWLEY, P. R. The importance of clinical psychology in a complete medical program. *J. consult. Psychol.*, 1946, **10**, 292-300.

59. HEALY, W., & BRONNER, AUGUSTA F. The child guidance clinic: birth and growth of an idea. In L. G. Lowrey (Ed.), *Orthopsychiatry, 1923-1948: retrospect and prospect*. New York: American Orthopsychiatric Association, 1948. Pp. 14-49.

60. HEISER, K. F. The need for legislation and the complexities of the problem. *Amer. Psychologist*, 1950, **5**, 104, 108.

61. HOBBS, N. The development of a code of ethical standards for psychology. *Amer. Psychologist*, 1948, **3**, 80-84.

62. HOCH, A. A review of psychological and physiological experiments done in connection with the study of mental diseases. *Psychol. Bull.*, 1904, **1**, 241-257.

63. HUNT, W. A. Clinical psychology—science or superstition. *Amer. Psychologist*, 1952, **6**, 683-688.

64. JAMES, H. (Ed.) *The letters of William James*. Vol. II. Boston: Atlantic Monthly Press, 1920.

65. JAMES, W. *The principles of psychology*. New York: Holt, 1890.

66. JAMES, W. *The varieties of religious experience*. New York: Longmans Green, 1902.

67. KELLY, G. A. Single level versus legislation for different levels of psychological training and experience. *Amer. Psychologist*, 1950, **5**, 109, 111.

68. KRAEPELIN, E. *Lehrbuch der psychiatrie*. Leipzig: Verlag von Johann Ambrosius Barth, 1899.

69. LONGSTAFF, H. P., SPEER, G. S., McTEER, W., & HARTSON, L. D. A survey of psychologists in four mid-western states. *Amer. Psychologist*, 1950, **5**, 422-423.

70. LOUTIT, C. M. The nature of clinical psychology. *Psychol. Bull.*, 1939, **36**, 361-389.

71. LOWREY, L. G. The birth of orthopsychiatry. In L. G. Lowrey (Ed.), *Orthopsychiatry, 1923-1948: retrospect and prospect*. New York: American Orthopsychiatric Association, 1948. Pp. 190-216.

72. MENNINGER, W. C. The relationship of clinical psychology and psychiatry. *Amer. Psychologist*, 1950, **5**, 3-15.

73. MERRILL, MAUDE A. Oscillation and progress in clinical psychology. *J. consult. Psychol.*, 1951, **15**, 281-289.

74. MEYER, A. G. Stanley Hall, Ph.D., LL.D. *Amer. J. Psychiat.*, 1924-25, **81**, 151-153.

75. MEYER, A. AUGUST HOCH, M.D. *Arch. Neurol. Psychiat.*, 1919, **2**, 573-576.

76. MEYER, A. Organization of community facilities for prevention, care, and treatment of nervous and mental dis-

eases. *Proc. First Inter. Cong. ment. Hyg.*, 1932, 1, 237-257.

77. MORRIS, L. *William James: the message of a modern mind*. New York: Scribners, 1950.

78. MORROW, W. R. The development of psychological internship training. *J. consult. Psychol.*, 1946, 10, 165-183.

79. MURCHISON, C. (Ed.) *The psychological register*. (2 vols.) Worcester: Clark Univer. Press, 1929, 1932.

80. MURPHY, G. *Historical introduction to modern psychology*. (Rev. Ed.) New York: Harcourt Brace, 1949.

81. MURRAY, H. A., et al. *Explorations in personality: a clinical and experimental study of fifty men of college age*. New York: Oxford Univer. Press, 1938.

82. NORSWORTHY, NAOMI. The psychology of mentally deficient children. *Arch. Psychol.*, N. Y., 1906, No. 1.

83. PETERSON, J. *Early conceptions and tests of intelligence*. Yonkers, N. Y.: World Book Co., 1925.

84. PINTNER, R. *Intelligence testing: methods and results*. (New Ed.) New York: Holt, 1931.

85. PINTNER, R., & PATERSON, D. G. *A scale of performance tests*. New York: Appleton-Century, 1917.

86. ROBACK, A. A. *History of American psychology*. New York: Library Publishers, 1952.

87. ROGERS, C. R. Where are we going in clinical psychology? *J. consult. Psychol.*, 1951, 15, 171-177.

88. ROSENZWEIG, S. Imbalance in clinical psychology. *Amer. Psychologist*, 1950, 5, 678-680.

89. ROSENZWEIG, S. Balance in clinical psychology: a symposium in correspondence. *Amer. Psychologist*, 1951, 6, 208-212.

90. SCOTT, W. D. An interpretation of the psycho-analytic method in psychotherapy with a report of a case so treated. *J. abnorm. Psychol.*, 1909, 3, 371-377.

91. SEASHORE, C. E. *Pioneering in psychology*. Iowa City, Iowa: Univer. of Iowa Press, 1942.

92. SHAKOW, D. One hundred years of American psychiatry: a special review. *Psychol. Bull.*, 1945, 42, 423-432.

93. SHAKOW, D. Clinical psychology: an evaluation. In L. G. Lowrey (Ed.), *Orthopsychiatry, 1923-1948: retrospect and prospect*. New York: American Orthopsychiatric Association, 1948. Pp. 231-247.

94. SHAKOW, D. Psychology and psychiatry: a dialogue. *Amer. J. Orthopsychiat.*, 1949, 19, 191-208, 381-396.

95. SHARP, STELLA E. Individual psychology: a study in psychological method. *Amer. J. Psychol.*, 1899, 10, 329-391.

96. SIDIS, B. Studies in psychopathology. *Boston med. & surg. J.*, 1907, 156, 321-326, 357-361, 394-398, 432-434, 472-478.

97. SMITH, T. L. The development of psychological clinics in the United States. *Ped. Sem.*, 1914, 21, 143-153.

98. SPEER, G. S. A survey of psychologists in Illinois. *Amer. Psychologist*, 1950, 5, 424-426.

99. STEVENSON, G. S., & SMITH, G. *Child guidance clinics: a quarter century of development*. New York: Commonwealth Fund, 1934.

100. SYMONDS, J. P. Ten years of journalism in psychology, 1937-1946; first decade of the Journal of Consulting Psychology. *J. consult. Psychol.*, 1946, 10, 335-374.

101. TERMAN, L. M. Trails to psychology. In C. Murchison (Ed.), *A history of psychology in autobiography*. Vol. II. Worcester: Clark Univer. Press, 1932. Pp. 297-332.

102. U. S. DEPT. ARMY, OFFICE OF THE SURGEON GENERAL. The U. S. Army's senior psychology student program. *Amer. Psychologist*, 1949, 4, 424-425.

103. VETERANS ADMINISTRATION. Cooperative training program for clinical psychologists in association with part-time work in VA stations where neuro-psychiatric cases are treated. *V. A. Tech. Bull.*, 1948, TB 10A-146.

104. VESTERMARK, S. D. Training and its support under the National Mental Health Act. *Amer. J. Psychiat.*, 1949, 106, 416-419.

105. WALLIN, J. E. W. *The mental health of the school child*. New Haven: Yale Univer. Press, 1914.

106. WATSON, R. I. The professional status of the clinical psychologist. In R. I. Watson (Ed.), *Readings in the clinical method in psychology*. New York: Harper, 1949. Pp. 29-48.

107. WATSON, R. I. *The clinical method in psychology*. New York: Harper, 1951.

108. WELLS, F. L. *Mental tests in clinical practice*. Yonkers, N. Y.: World Book Co., 1927.

109. WENDT, G. R. Legislation for the general practice of psychology versus legislation for specialties within psychology. *Amer. Psychologist*, 1950, **5**, 107-108.
110. WHIPPLE, G. M. *Manual of physical and mental tests*. Baltimore: Warwick and York, 1910.
111. WHITEHORN, J. C. A century of psychiatric research in America. In J. K. Hall (Ed.), *One hundred years of American psychiatry*. New York: Columbia Univer. Press, 1944. Pp. 167-193.
112. WIENER, D. N. The Minnesota law to certify psychologists. *Amer. Psychologist*, 1951, **6**, 549-553.
113. WISSLER, C. The correlation of mental and physical tests. *Psychol. Monogr.*, 1901, **3**, No. 6 (Whole No. 16).
114. WITMER, L. Clinical psychology. *Psychol. Clin.*, 1907, **1**, 1-9.
115. WITTY, P. S., & THEMAN, VIOLA. The psycho-educational clinic. *J. appl. Psychol.*, 1934, **18**, 369-392.
116. WOLFLE, D. The reorganized American Psychological Association. *Amer. Psychologist*, 1946, **1**, 3-6.
117. WOLFLE, D. Legal control of psychological practice. *Amer. Psychologist*, 1950, **5**, 651-655.
118. WOODWORTH, R. S. *Personal Data Sheet*. Chicago: C. H. Stoelting, 1917.
119. WYATT, F. Clinical psychology and orthopsychiatry. In L. G. Lowrey (Ed.), *Orthopsychiatry, 1923-1948: retrospect and prospect*. New York: American Orthopsychiatric Association, 1948. Pp. 217-230.
120. YERKES, R. M. (Ed.) *Psychological examining in the United States Army*. (Memoirs of the National Academy of Sciences, Vol. 15.) Washington, D. C.: U. S. Govt. Printing Office, 1921.
121. YOUNG, K. The history of mental testing. *Ped. Sem.*, 1924, **31**, 1-48.

Received January 19, 1953.

## PSYCHOLOGY IN ITALY

HENRYK MISIAK  
*Fordham University*

AND

VIRGINIA M. STAUDT  
*Hunter College*

Italy very early took cognizance of scientific psychology and joined other countries in accepting and developing the new science (2, 6, 9, 28). Although she made significant contributions to the study of human behavior, her influence in the history of psychology was never very profound.

The attainments of Italian psychologists have not been equal in importance to those made by Italian scientists in allied areas like psychiatry and neurology. The contributions of Italy in these latter fields have attracted worldwide attention (16). The first homes for the insane were established in Italy as early as the 14th century, and the psychiatric hospital Santa Maria della Pietà in Rome, founded in 1548 and still functioning, is the oldest in the country. The reforms in the treatment of patients, effected by Chiarugi (1759-1820), actually antedated those of the French physician Pinel (1745-1826), who is credited with being the first to revolutionize the treatment of the insane. Chiarugi was also the author of the first Italian text in psychiatry. Music therapy, occupational therapy, and even psychodrama were used by Miraglia in Naples in the middle of the 19th century. In Italy psychiatry has always had a neurological orientation. Therefore, it is not surprising that in recent times in Italy electroshock (Cerletti, Bini) and acetylcholine induced shock (Fiamberti) should have originated. Supraorbital lobotomy in psychosurgery (Fiamberti) also had its origins in Italy. In

neuroanatomy and neurology such names as Golgi, Luciani, Marchi, Bianchi and their discoveries are well known.

### BEGINNINGS AND PIONEERS

Materialistic positivism had a strong hold on early Italian psychology mainly because of the powerful and enduring influence of Roberto Ardigò, professor at Padua University, who in 1870 published *La Psicologia come Scienza Positiva* and in 1898 *Unità della Coscienza* in which he identified all mental life with cerebral physiology. In the idealistic reaction against positivism, psychology, which was considered an offspring of this system and a pseudophilosophy, was opposed together with positivism. As a result psychology was excluded from, or greatly reduced in, the school curricula. The early strongholds of psychology were the universities at Rome, Florence, and Turin.

The first Italian psychologist was Giuseppe Sergi (1841-1936), professor at the University of Rome. The year of publication of his book *Principi di Psicologia*, 1873, the same year as that of Wundt's first volume of *Grundzüge der physiologischen Psychologie*, is regarded as the birth date of Italian psychology. In 1885 the first Italian psychological laboratory was established by him as a section of the Institute of Anthropology in Rome. Although Sergi wanted psychology to be independent of philosophy, his own philosophical credo, strong positivism, permeated his views in psychology.

A different philosophy and a different philosophical trend were represented by another pioneer of psychology in Italy, Francesco de Sarlo (1864-1937), a devoted disciple of Brentano, at the University of Florence (4). Through his efforts the first institute of psychology was opened in Florence in 1903. De Sarlo was a philosopher, psychiatrist, surgeon, and psychologist. His views in psychiatry anticipated the psychobiological concept, later to become so popular, but without violation of the traditional dualism. He received his training from the director of the Psychiatric Hospital in Reggio Emilia, Augusto Tamburini, a man who was greatly interested in psychology and influenced its development in Italy through his students, such as De Sarlo, G. Buccola, and C. G. Ferrari. It was Tamburini who first popularized the concept of mental hygiene among the Italians. Buccola (1854-1885) was the author of *La Legge del Tempo nei Fenomeni del Pensiero* (1883), a work which aroused a current of sympathy for psychological research in Italy. Ferrari (1869-1932) contributed substantially to the development of psychology: with Tamburini he founded a psychological laboratory in Reggio Emilia in 1896, the magazine *Rivista di Psicologia*, and in 1901 he translated W. James's *Principles of Psychology*, an event which has been acknowledged as the turning point in the history of Italian psychology (9, 10).

In Turin psychological problems were introduced into the field of physiology by Angelo Mosso (1846-1910) and into criminal anthropology by Cesare Lombroso (1835-1909). The former is well known to psychologists for his ergograph and for his pioneering research on work and

fatigue. His books, *La Paura* (1884) and *La Fatica* (1891), were translated into many languages, including English. Mosso became famous for his *L'Uomo Delinquente* (1876) and was hailed as the forerunner of constitutional psychology. In general, Italians, and especially such scientists as Di Giovanni, Viola, and Pende, have contributed to constitutional psychology by their anthropometric techniques and their vigorous efforts to establish the relationship between physique and personality. Although Lombroso did not develop any system, his observations and his views aroused a great deal of controversy and prompted much research. In later years he directed his attention to parapsychology. In 1895 Mosso's laboratory was turned over to Friedrich Kiesow (15) (1858-1940), a pupil and former assistant of Wundt's and a friend of Külpe's. Kiesow transplanted from Leipzig to Turin not only the knowledge of experimental techniques, but also an enthusiasm for experimentation in psychology. For many years, long before others opened laboratories and began research in other places, he was the greatest experimentalist in Italy. His own work was mainly in the field of sensation, in taste and touch especially. Among Kiesow's students were Gemelli and Ponzo. Another experimentalist imported from abroad was Vittorio Benussi (1878-1927), who, although born in Italy, lived in Austria for many years and became the most outstanding experimentalist there. After World War I he went to Padua, taught psychology at the university there, and opened a laboratory. His studies on perception, respiration, suggestibility, and hypnosis received recognition in Italy and abroad.

A special place among Italy's

pioneers in psychology is held by Sante de Sanctis (1863-1935), an eminent figure not only in psychology, but also in psychiatry (1, 7). Through his enthusiasm for the new science, his versatile activities, his writings, and his personal charm he left a deep impression on Italian psychology. He succeeded his teacher, Sergi, as head of the school of Rome and became the first graduate teacher of experimental psychology in Italy. While practically all the fields of theoretical and applied psychology were explored by De Sanctis, his best energies were devoted to child study on the one hand and to the understanding and helping of the mentally deficient and the abnormal on the other hand. His publications were numerous, original, and influential. Interested in sleep and dreams, De Sanctis published several studies on the subject in 1896 and a large monograph, *I Sogni* (Dreams), in 1899 which preceded that of Freud. Other important writings include *La Mimica del Pensiero* (1904), *Psicologia Sperimentale* (Vol. I, 1929; Vol. II, 1930), and *La Conversione Religiosa* (1924) (translated into English, *Religious Conversions: a Bio-Psychological Study*, [1927]). He professed recognition of and respect for religious faith. "I had dealings with materialists, positivists, rationalists," he said, "but they all failed to inculcate into me their philosophical convictions," and he said further that no one succeeded in shaking his belief in finality in nature and the universe (7).

#### PROGRESS AFTER 1905

In 1905 a significant event took place in Italy, the 5th International Congress of Psychology in Rome, under the presidency of Sergi, one of the most provocative congresses of

psychology ever held. At this congress the first scale of intelligence, the Binet-Simon, and W. James's paper, "La Conscience Existe-t-Elle?" were given prominence. In general the congress found psychology in Italy firmly established. The work of Italian psychologists and of their laboratories was already well recognized. There was a lively interest in child study. The new studies in this field conducted in Britain, Germany, and France were matched by the excellent medical and biological studies of the child in Italy. Melzi, De Sanctis, and Ferrari had already distinguished themselves in that field, and Maria Montessori was about to launch her new educational movement. There were two dominant currents of psychological thought, one stemming from Wundt, the other from Münsterberg (6).

There was one significant practical outcome of the congress for Italy: the next year the ministry of education formally instituted three autonomous chairs of psychology on the university level, in Rome, Turin, and Naples, given respectively to De Sanctis, Kiesow, and Colucci. De Sanctis was later succeeded by Mario Ponzo, Kiesow by Alessandro Gatti, and the latter after his premature death by Angiola Massucco Costa. Galdo succeeded Colucci.

The other centers of psychological research and teaching have been Florence, Padua, Milan, and Genoa. In Florence the institute of psychology was headed by De Sarlo, then by Enzo Bonaventura, and since 1938 by Alberto Marzi. The chair of psychology in Padua was occupied by Benussi, and from 1927 to 1948 by Cesare Musatti. In 1948 when Musatti accepted the chair of psychology at the State University in Milan (founded three years after the

Catholic University in Milan, in 1924) he was succeeded at Padua by Fabio Metelli. Musatti, who has been one of the few exponents of psychoanalysis, is the author of a two-volume work on the subject. In the other university of Milan, the Catholic University of the Sacred Heart, Agostino Gemelli has been the head of the psychology department. In Genoa Giuseppe Vibone has taught psychology at the University. There the faculty of medicine has also instituted a chair of psychology, of which Amedeo Dalla Volta is head.

In World War I, during the political upheaval in Italy in the nineteen thirties, and more recently in World War II, psychology suffered a great blow. In spite of these hindrances progress has continued, however, marked among other events by the creation of the Italian Psychological Association and by the state decree in 1935, making psychology a required subject for a degree in philosophy and complementary for a degree in pedagogy, medicine, and law (19, 20, 27). The opening in 1940 of an experimental center for applied psychology in the National Research Council, with Ferruccio Banissoni as director, is to be mentioned particularly for it gave a new impetus to psychological research (3).

#### TRENDS IN ITALIAN PSYCHOLOGY

Following the period of interest in psychophysics, perception, and child study, so prominent prior to 1920, the emphasis in contemporary Italian psychology shifted to applied fields. A number of private and public centers for psychotechnical research were opened and almost all leading psychologists were engaged at some time or other in applied psychology. Aptitude testing and vocational guidance received special prominence. There

were strong efforts to extend psychology to education through the development of new methods of education based on the findings of child psychology, by instituting projects for school reform, by training teachers in psychology, and by similar means. Child psychology has been fostered at the University of Florence. The psychology of personality, or characterology as it is called there, has attracted many adherents in Italy. Psychoanalysis, however, as a doctrine and method has had very few followers and is taught at very few universities. The religious, cultural, and social traditions of Italy have not allowed psychoanalysis to enjoy the favorable position which it has had in other countries; in fact, there have always been more opponents than defenders of psychoanalysis (23).

#### LEADING CONTEMPORARY SCHOOLS AND PSYCHOLOGISTS

Let us now review the important centers of psychology and acquaint ourselves with the leading psychologists of Italy. It is obvious that it is the ability and talents of the men at these universities that determine the role and influence of these schools on Italian psychology. Thus it will be mainly the psychologists themselves on whom we shall focus our attention. Following a historical sequence, we shall speak of Ponzo and the University of Rome, Marzi and Metelli and the University of Florence, Bari and Padua, Gemelli and the Catholic University of Milan, and Banissoni and the National Institute of Psychology. Then we shall mention the new centers of psychology at the Gregorian University and at the Salesian University. The most prominent living psychologists from the point of view of the volume and

extent of work and influence are Gemelli, Ponzo, Marzi, and Metelli.

#### *University of Rome*

As mentioned before, the teaching and the laboratory of psychology were begun here by Sergi. Under the subsequent direction of Sante de Sanctis, student and collaborator of Sergi, the school of Rome flourished and gained the prestige which it still has under the leadership of Mario Ponzo, successor of De Sanctis. Ponzo has become the most representative psychologist of Italy. Born in 1882 in Milan, he studied in Turin, where he received the doctorate in medicine, and then taught psychology at the University of Turin from 1905 to 1931. He was Kiesow's pupil and then his colleague and collaborator. In 1931 he went to Rome and has been there ever since as a professor of psychology and head of the Institute of Psychology. Ponzo is the president of the Italian Psychological Association, which after many years of inactivity was revived in 1951. He is also a member of the Executive Committee of the International Union of Scientific Psychology. An indefatigable worker and an unusually prolific writer, Ponzo has for a long time been associated with psycho-physical problems, especially touch and taste. Later he extended his research to other fields of psychology and in the last two decades he has devoted himself particularly to applied psychology. With a profound conviction in the usefulness and possibilities of applied psychology, Ponzo has strenuously fought for its recognition in education, industry, vocational guidance, and personnel selection. His own studies are valuable contributions to aptitude testing and professional selection. According to Ponzo, vocational guidance should

not be based solely on the assessment of aptitudes or intelligence, but on the general personality characteristics. Personality, character, and the natural inclinations of the individual were frequently stressed by Ponzo as the primary consideration, whether in vocational guidance, personnel selection, or accident prevention. The recognition and growth of applied psychology and the emphasis on the problems of applied psychology in contemporary Italian psychology were undoubtedly to a great degree the result of Ponzo's writings and activities, which also produced practical effects in various spheres of the national life of Italy.

Upon succeeding De Sanctis, Mario Ponzo dedicated himself to organizing the Institute of Psychology of the University of Rome, which he unfortunately saw ruined during the war as a result of the bombings. Since World War II, owing to his efforts and zeal, the Institute has again revived its activity and offers much promise of promoting psychological studies. Each member of the staff has his particular area of interest. Ponzo himself at present is interested in group psychology and vocational guidance. Leandro Canestralli, assistant director at the Institute, works primarily in the areas of measurement of psychomotor skills and general experimental psychology and has distinguished himself by original studies of psychomotor activity and the use of the photocyclographic method photographing the trajectories covered during movement in the study of voluntary acts.

The second assistant, Ernesto Valentini, is an animal psychologist. Gigliola Sbordoni specializes in the field of psychometrics, child psychology, and audiovisual aids in teach-

ing. The remaining three assistants serve without compensation. Franco, Ferracuti, and Sbrana work in clinical psychology, Franco being especially interested in the psychological problems of senescence and of cancer patients. A clinical psychologist in the Italian Air Force, Nicolo Numeroso, is working on personnel selection and is also interested in psychodrama. In addition to their experimental work, Ponzo and his staff also teach rather large psychology classes, conduct a child guidance clinic and a school for social workers, have a private clinical practice, and do some industrial counseling.

The laboratory at Rome itself is well equipped as far as the older standard apparatus and space are concerned. Experienced machinists build special apparatus for the experimentation and there are excellent facilities for constructing equipment. Most of the research is done primarily in the field of perception. An account of the work accomplished during the last five years in the laboratory is presented in the Proceedings of the IXth Congress of Italian Psychology (Nov. 1-4, 1951). There are lecture halls, which are very suitable for laboratory demonstrations, and visual aids including movie projection equipment. The library is well provided with Italian, French, and German journals, but the American and English journals are noticeably lacking.

#### *University of Florence*

De Sarlo, founder and director of this first institute of psychology in Italy, was succeeded by Enzo Bonaventura, who demonstrated the significance of empirical factors in the genesis of perception by his research on perception, especially on the perception of time and space. Bonaventura published several volumes; the most successful was his work on psychoanalysis, *La Psicoanalisi* (last edition in 1950). Because of anti-Semitic attitudes in Italy at that time he was forced to relinquish his post as director of the institute and was finally compelled to leave Italy. Up to the time of his tragic death in 1948, which occurred during an Arab attack when he was giving aid during the war in Palestine, Bonaventura held the seat of psychology at the University of Jerusalem.

Bonaventura's successor in 1938 was Alberto Marzi, a prominent figure of growing influence in Italian psychology today. His research has included a variety of fields such as attention, eidetic imagery, the nature of intelligence of deafmutes, and psychotechnical problems. Marzi was associate professor at the University of Florence until 1948 and then after taking the requisite government examinations he became professor. Having received a Rockefeller grant, he visited the United States for five months. Following this visit he proceeded to the National Institute of Psychology in London and then to Paris where he studied child psychology with Henri Wallon. When Marzi returned to Italy he became the head of psychology at the University of Bari in southern Italy, the fourth most important Italian university. Thus Marzi continues to hold academic appointments at Florence and at Bari, a fact that indicates the shortage of senior psychologists in Italy (14). There are also other psychologists with dual appointments, such as Banissoni and Metelli (11). The modest Laboratory of Industrial Psychology in Florence, directed by Marzi, is financed by the city and it provides testing services to industry and to the school system.

Marzi is assisted by Maria Louisa Falorni, who directs the practical work of four young psychologists and seven social workers. His other activities include the direction of the Center for Studies of Accidents which was established in 1951 and which has a staff of five psychologists, all graduates of the University of Florence, and one social worker.

In addition to Marzi's teaching duties in Florence he has been translating the Wechsler-Bellevue test for children with the assistance of Lydia de Rita and he has been studying projective techniques with Giovanni Battista Guarini. At Bari, however, Marzi's major interest has been the psychological study of social classes. He has been engaging in a study of some seven hundred cave dwellers from the nearby villages of Matera and Lucania.

Despite his many teaching and research interests and his travels between Florence and Bari, Marzi finds time to edit *Rivista di Psicologia*, the bulletin of the Italian Psychological Association, of which he is general secretary. He is also a foreign affiliate of the American Psychological Association. Marzi's work at Florence has been published in two volumes entitled *Studie Ricerche Università delle Firenze 1938-1947* (Vol. I) and *1947-1949* (Vol. II). He has prepared a third volume on *Studie Ricerche Istituto di Bari* (14). At present he is writing a book on the psychology of work and he plans to translate the next edition of Viteles' *Industrial Psychology*.

Fabio Metelli presides over the laboratory which is operated under the auspices of the department of philosophy of the University of Florence, where he is assistant professor. Strangely enough, however, Metelli's real teaching appointment is

at the University of Padua, and he spends three days each week in Florence and three at Padua. Metelli considers the continuation of the laboratory at Florence a distinct service to psychology (11). He attributes the lack of adequate support for the laboratory, which is operated on an annual budget of about eighty dollars, to the fact that Italian philosophers are too idealistic and therefore are not very enthusiastic about science. His own research and writings have covered such areas as dreams, testimony, and the perception of movement.

At the University of Padua one course is given in psychology which varies each year—individual differences, memory, perception. Consequently, in order to cover the field the student must continue to take this course for several years. He must also get a good deal of his background in psychology from his own reading. Demonstrations are provided in the course, but there is no opportunity for laboratory practice. The laboratory at Padua, located in a 15th century building, has very limited space, apparatus, and library facilities. Much of the research done is in the field of configural perception after the tradition of Benussi (11).

#### *Catholic University in Milan*

The best equipped laboratory, not only in Italy but probably in all continental Europe, is at the Catholic University in Milan (14). It was established and has been directed by a Catholic priest, Agostino Gemelli, the most prominent contemporary Italian psychologist, whose influence in this Catholic country (99.6% of the inhabitants are Catholic) has been very significant with respect to the attitude of Catholics towards

psychology. This influence marks a new era in Italian psychology.

Gemelli (5, 12, 18, 20) was born in 1878 and he studied medicine at the University of Pavia. After his doctorate in medicine and surgery he remained at the university and continued research as an assistant to Professor Camillo Golgi (1844-1926), famous anatomist and physician, Nobel prize winner. During this period he also studied philosophy and was engaged in social and political activities. Abandoning religion, he embraced materialism and Marxism, but gradually became dissatisfied and disillusioned with these systems. Finally he returned to [the Church, entered the Franciscan order, and was ordained a priest in 1906. From 1907 to 1911 Gemelli studied biology, physiology, and philosophy at various universities in Bonn, Frankfort on Main, Munich, Cologne, Vienna, Louvain, Amsterdam, and Paris. In 1911 he completed his doctorate in philosophy at the University of Louvain.

The psychologist who gave Gemelli his initial training in experimental psychology was Friedrich Kiesow, who from 1899 was on the faculty of the University of Turin and who recommended Gemelli to his friend Oswald Külpe. Gemelli went to Külpe at Bonn and then with him to Munich, working in his laboratories. Gemelli has always recognized the profound impression that Külpe made on him, and through the years he has continued to regard himself as a "student of Külpe" (12). The interrupted association with Kiesow was renewed again and continued afterwards for many years when Gemelli accepted the appointment by the Italian government to teach psychology at the University of Turin in 1914. With Kiesow he

founded the *Archivio di Psicologia, Neurologia e Psichiatria*.

During World War I Gemelli served as a military chaplain, a physician, and a psychologist. In the latter capacity he became known for his work in the selection of pilots. After the war his efforts were directed towards the founding of a Catholic university which was officially opened in 1921 in Milan under the name "The Catholic University of the Sacred Heart in Milan." From its founding Gemelli has been the rector of the university and has done a great deal to promote the teaching of psychology and psychological research. In the subsequent years Gemelli has conducted a large number of original researches and published many books and articles. His bibliography comprises several hundred items, among them such books as *La Psicologia del Pilota di Velivolo* (1942), *La Psicologia Applicata all'industria* (1944), *L'Orientamento Professionale dei Giovani nelle Scuole* (2nd Rev. Ed., 1947), *Introduzione alla Psicologia* (2nd Rev. and enlarged Ed., 1949). These are simply the latest works.

From a perusal of Gemelli's writings, taken in chronological order, it is evident that his initial studies were in the most difficult area of psychology, the so-called higher mental processes, thought and will, an obvious outcome of his association with Külpe and Michotte. Then when he was with Kiesow in Turin the major areas of his work were psychophysics, sensation, and perception. At Milan in his own laboratory the most original was his research on language. In the meantime his interest in applied psychology grew, a field to which he made appreciable contributions, particularly in aviation, industry, education, and criminology.

As a spokesman for, and interpreter of, the new psychology for Catholics, Gemelli was forceful and effective. He clarified for Catholics the status of psychology as a science, defined its relationship to philosophy, biology, and other disciplines, and most of all devoted himself to experimental research, contributing to psychology, adding something of his own to its treasure. He dispelled the last shreds of suspicion and doubt about the new psychology, especially in Italy. His authority and prestige carried weight. After all, he was a physician, a psychologist, a theologian, a priest, and a monk all in one. Moreover, he was a rector of a Catholic university and the president of the Pontifical Academy of Sciences. His own example spoke louder than his words. That there can be harmony between science and faith, between modern progress and old tradition, between psychology and philosophy, and between speculation and experiment, his writings explained and his example proved.

Biological orientation has been a characteristic feature of Gemelli's psychology. Some Catholic psychologists have the tendency, perhaps, to let the biological counterpart of man be overshadowed by the spiritual or mental, to divide man too much into the physical and psychical, to minimize the bodily component in psychological processes. Gemelli has avoided that. Whether in emotion, or perception, or delinquency, or in other fields, he has been fully aware of organismic forces. The biological orientation which also made him bring psychology closer to biology than to philosophy resulted from both his medical training and his consistent application of the Aristotelian hylomorphic doctrine of man to his work and theories. What-

ever aspect of man he considered or examined, he never lost sight of the unity of man.

In the Catholic University of the Sacred Heart in Milan psychology received special attention. The systematic teaching of experimental psychology was included in the program of the university and a special laboratory for training and psychological research was opened. Gemelli assumed the role of professor of psychology and director of the laboratory (18).

The work in the Laboratory of Psychology has been very fruitful. Its productivity is demonstrated in the numerous volumes of a special series of monographs, *Contributi del Laboratorio di Psicologia*, which contains original studies completed by Gemelli and his collaborators. The areas of research at the Laboratory are varied, but the best studies have been done in the field of perception, personality, and language. In perception the specific problems investigated included: individual differences, localization of sound, perception of depth, and spatial perception. The theoretical stand assumed on the basis of these studies places the school of Milan between the elementaristic view along the lines of Wundtian tradition and that of the Gestalt school. The school maintains that even though we normally have a unification of sensory data, there is also the possibility of perceiving isolated elements without constituting them into a whole. In the field of motor activity with human and animal subjects it was demonstrated how a task unifies and renders intelligible various phases of activity. The most original and brilliant was the study of language based on the electroacoustic recordings which gave new insight into the

understanding of language. Extensive research was done in electro-encephalography, the Laboratory being the first to begin this type of research in Italy. During World War II the Laboratory engaged in special work for both the war effort and the city of Milan. Other areas include psychometrics, child psychology, social psychology, industrial psychology, and gerontology.

During the war the buildings of the Laboratory were completely destroyed but the equipment was saved. The total reconstruction took several years and only in 1950 could the Laboratory resume its work fully. Gemelli now has nine assistants. Five of them have a doctorate in medicine, three in pedagogy, and one in philosophy. There are about twenty graduate students working on different problems for their dissertations.

A prominent figure at the University of the Sacred Heart has been a priest, Giorgio Zunini (born in 1903), professor of psychology there and Gemelli's closest collaborator. He began his scientific career at the University of Pavia. From geology and paleontology he turned to biology and then to comparative psychology and systematic psychology. Among his publications in comparative psychology are excellent and original studies on learning in fish. In his book, *Animali e Uomo; Visti da uno Psicologo* (1947), he shows the gradual development of behavior from fish to man and the qualitative differences existing between man and animals. The systematic works include *Introduzione alla Psicologia* (2nd Ed., 1949) published with Gemelli and *Psicologia* (2nd Ed., 1950), a brief history of psychology and exposition of psychological systems. Both books filled a need that

was sorely felt in Italian psychological literature.

The University of the Sacred Heart has been the place where many psychologists have begun their professional careers. A number of Italians as well as foreigners received their training and degrees there and carried out their studies using the facilities of the Psychological Laboratory. Among the Italians there were several who distinguished themselves. Arcangelo Galli, a pupil of Michotte at Louvain and of Wertheimer and Gell at Frankfort, did several experimental studies at Louvain (1914, 1924). Even before the founding of the Catholic University he published a study with Gemelli and Tessier on the perception of the position of the body. During the many years of his association with the Laboratory he made several contributions mainly in the field of perception as well as some in applied psychology in collaboration with Gemelli. Alessandro Gatti (25), author of articles on psycho-technical and perceptual problems and of a book, *Le Massime e i Caratteri* (1934), went to the University of Turin, where he founded a center of studies on work, the only center of this type in Italy. In 1928-1929 Gatti visited the United States as a Laura Rockefeller fellow. He died in Turin at the height of his professional activity. Giuseppina Pastori, now director of the Laboratory of Biology at the Catholic University, collaborated with Gemelli on his electroacoustic study of language. G. Sacerdote became director of the Institute of Electroacoustics in Rome. C. Trabattoni worked in electro-encephalography, and L. Ancona studied stereoscopic "aftereffects."

The students at the University and investigators at the Laboratory have come from many foreign countries:

from Austria, Hubert Rohracher, now professor at the University of Vienna and the most prominent experimentalist in Austria; from Bulgaria, B. Raduseff; from Finland, A. Penttila; from Yugoslavia, A. Terstenjak, P. Gubernia, P. Matko; from Lithuania, A. Sidlauskas, now assistant professor at the University of Ottawa, Canada; from Rumania, A. Manoil, who published a large volume on the Milan School (18), and who is now chairman of the department of psychology at Park College, Parkville, Missouri; from China, P. Siao Sci Wi. It is evident from this enumeration that the influence of the Catholic University in Milan has been very widespread.

#### *The National Institute of Psychology in Rome*

The National Institute of Psychology had its origins in the Experimental Center of Applied Psychology of the National Research Council in 1940. Later in 1946 it was called the Center of Psychological Studies and then on January 1, 1950 its title was changed to the National Institute of Psychology. There are fourteen laboratories affiliated with the Institute, and of this number ten are situated in the university cities of Bari, Bologna, Catania, Florence, Genoa, Messina, Naples, Padua, Trieste, and Turin.

The Institute and its collaborating laboratories attempt to furnish scientific counsel and guidance to public and governmental offices on a variety of psychological problems, such as the selection of scientific and civil service personnel, air force, military, naval, and police personnel selection, vocational guidance for students at the age of eleven years for classical and higher education, industrial counseling, and accident prevention. As a

result of these studies and of the collection of a tremendous amount of data, it is possible to make population analyses and comparative studies of the inhabitants of the various regions of Italy.

The psychological activities of the Institute have also extended to other areas. New tests are being devised; national standardizations are being developed; foreign tests, such as the Terman-Merrill, are being translated, adapted, and standardized. A variety of research investigations are being conducted on such topics as the interrelationships of personality and environment, characterology, and prejudice. In 1951 at the 13th International Congress of Psychology in Stockholm and at the 10th International Congress of Psychotechnics at Gothenberg, Banissoni gave an account of some of the Institute's psychosocial studies. Through the Institute's research, Professor Banissoni hoped to emphasize the usefulness of applied psychology and thus to encourage the universities to foster and improve the teaching of general and experimental psychology which for years had been poorly done in the departments of philosophy. The Center of Technico-Scientific Documentation of the National Research Council receives the reports of the National Institute of Psychology in the fields of general and applied psychology and publishes notices of them in its *Index of Scientific and Technical Periodicals*.

In addition to all these activities the director and staff members of the Institute offer training courses for the Institute's personnel and they also give courses to the Institute's external collaborators. For the most part, these courses are brief, conducted over a period from a few weeks to three months and covering a fairly

wide range of topics: statistics in psychological research, educational psychology, social psychology, applications of the Szondi projective test, and psychosomatic medicine. Such lectures and courses have been held in Florence, Trieste, Civitavecchia, Udine, and Salerno, as well as in Rome.

The organizer of the National Institute of Psychology was Ferruccio Banissoni. Born in Trieste in 1888, he studied in Vienna for four years and in Rome for two years. In 1921 he received the doctorate in medicine. For many years he was a professor at the University of Rome and he also lectured and gave courses at other institutions. His work, broad in scope and marked by originality and progressiveness, contributed to many fields of experimental, applied, and clinical psychology. With his training and research in medicine he was able to treat with authority the problems that were on the borderline of psychology and medicine, as for example his work on the application of the electrocardiogram in experimental psychology. Together with Ach, Michotte, De Sanctis, Lindworsky, and Abramowski, Banissoni is one of the few psychologists who studied experimentally one of the most difficult problems of psychology, the will, and wrote on the subject (*Contributo alla Psicologia Sperimentale della Volontà*, 1926). He and an outstanding Polish psychologist, M. Dybowski, had developed a plan for research on will problems in 1938, a plan which the latter carried out later in Poland (8). As head of the psychological section of the National Research Council, Banissoni inspired and directed a great many studies in applied psychology. The psychology of work, aptitude testing, personnel selection,

and a new branch of applied psychology called filmology have been the areas of his most fruitful research. In contemporary Italy it was Banissoni who raised psychology to new heights, and so his death in 1952 was a real loss to psychology.

#### *The Gregorian University in Rome*

For several centuries this venerable institution has been a place where clerical students from all over the world received their philosophical and theological training. The school is administered by the Jesuit Fathers and all the professors are Jesuits from various countries throughout the world. Since the University is primarily devoted to teaching seminarians and priests, it emphasizes philosophy and theology in its curriculum. Thus far Gregorian University has not offered many courses in scientific, social, or pastoral psychology and it has had no facilities for laboratory work, in spite of the presence of such excellent teachers of psychology as Joseph Fröbes, Johannes Lindworsky, Alexander Willwoll, and Paul Siwek. In recent years, however, a more serious effort has been made toward a better integrated program in psychology by a new professor, André Godin (born in 1915, Belgium). After working in criminal psychopathology under the direction of Professor De Greeff (Louvain), he received his doctorate in philosophy in Brussels, Belgium (1942), where he later stayed at the "Centre de Consultations Médico-Psychologiques," thus acquiring both good theoretical foundations and extensive practice in counseling and guidance. In 1949 he was named assistant professor of educational psychology at the Gregorian University. Feeling the need for additional

knowledge of experimental methods and statistics in psychology, he came to the United States and studied at Fordham University where he received a master's degree in psychology. During this period of study he visited many centers of psychological research in America. Upon returning to Rome he took with him not only a thorough knowledge of American psychology, but in addition some of the modern psychological apparatus, psychological films, and books for his work at the university there. Besides publishing articles in various magazines here and abroad he has delivered papers at two international congresses of psychology: in 1948 at Edinburgh and in 1951 at Stockholm. Godin's main interest lies in clinical psychology, particularly in psychoanalytic doctrine and method. The major task that he has posed for his scientific work is to bridge the psychoanalytic dynamics with moral and religious values. In the summer of 1951 Godin helped to organize an international conference of Jesuit psychologists, the first of its kind, which met in England.

#### *The Institute of Experimental Psychology of the Pontifical Salesian University in Turin*

The creation of this Institute by the Salesian Fathers in 1938 was a significant stride in the progress of experimental psychology among Catholics in Italy (17). The Institute has classrooms, a library, and a spacious laboratory that is well equipped with modern apparatus and is recognized as one of the best in the country. The courses offered cover a wide range of psychology: general experimental psychology, genetic psychology, characterology and biotypology, social psychology, psychological testing, ab-

normal psychology, industrial psychology, criminal psychology, and psychology of religion as preparation for teachers of religion. Seminars are also held. The main purpose of the Institute is to offer the latest findings in all fields of experimental psychology to the students who are engaged in philosophical studies. In original research the results of the Institute have not been impressive, but some interesting studies have been done with respect to vocational guidance and aptitude testing. During the war the Institute concerned itself with the psychological effects brought about by shock. Wounded soldiers and civilians who had been or were suffering from shock due to the bombing were studied and a program of rehabilitation for them was devised. In 1942 the Institute itself was bombed.

The director of this Institute has been a Salesian priest, Giacomo Lorenzini (born in 1909). Before becoming the director of the Institute of Experimental Psychology he taught philosophy at the Salesian Philosophical Institute of St. John Bosco in Turin. At the former Institute he has given courses in experimental and educational psychology. His main interest is in the psychology of adolescence with particular reference to its religious, social, moral, and esthetic aspects. Several of Lorenzini's publications deal with educational problems and personality typology. One is a systematic textbook, *Corso di Psicologia* (1948).

#### EVALUATION

In comparison to other countries psychology in Italy did not develop quickly. Experimental psychology was not as readily incorporated into the programs of institutions of higher

education. The appreciation of psychology and its influence on cultural and social life and on education have been very limited. At the meeting of the International Bureau of Education in Geneva in 1937, Italy was the only nation out of forty-two participating nations which did not agree to include psychology as a required course in the training of teachers. The chairs of psychology were instituted late and remained few. The number of psychological laboratories was small. In general, Catholic schools and institutions kept aloof. There was a strong resistance to psychology as an independent science. The recognition in wider circles of experimental psychology as a separate science, and not simply as a part of or as an adjunct to philosophy or physiology, came much later, only in the 1920's.

The reasons for the slow development of psychology and its limited influence in Italy have been multiple. They are of an economic, social, religious, and even a political nature. The proper evaluation of all these factors is not a simple task, but it is immediately obvious that the chief difficulty for psychology arose at the very outset when it was linked with a positivistic philosophy. Psychology used to be regarded as a product of that philosophy, and was practically identified with it. Thus the succeeding philosophical current, antagonistic to positivism and seeking to banish it, also rejected experimental psychology, its alleged product. Moreover the protagonists of the new psychology were often professed agnostics or outright materialists, not infrequently open enemies of religion who used psychology in their fight against religion. This state of affairs could not create a favorable atmosphere for psychology in Catholic

Italy, and instead of finding favor psychology met with opposition, suspicion, or indifference. The picture was altered only when a Franciscan priest, Agostino Gemelli, showed that experimental psychology is a separate autonomous science and *per se* does not contradict any philosophical or religious principles. The founding of the psychological laboratory at the Catholic University in Milan and the prestige of Gemelli overcame the suspicion and opposition toward experimental psychology among Catholics. The role of Gemelli and of the Catholic University in Milan in gaining acceptance for psychology by Catholics cannot be emphasized strongly enough. But one must not overlook the important influence of those Catholic psychologists whose prestige in Italy has been very great, namely Ponzo and Banissoni.

In spite of all these gains, psychology in Italy still has not attained the place and role that it should enjoy. It is still relegated for the most part to the faculty of philosophy as a complementary discipline. It attracts a very small following in university teaching. There are very few individuals who dedicate themselves to psychology because there is little possibility of remunerative work for a qualified psychologist. For the student who writes his thesis in psychology usually the only occupational prospect available is teaching philosophy (19, 20). Nevertheless psychology in Italy now is slowly but surely acquiring the scientific dignity and social prominence that it merits and has gained in other countries. As for the Catholic scientists and the Catholic institutions of learning, they are participating more and more in Italian psychology and their contributions are steadily growing in volume and quality.

## REFERENCES

1. APPICCIANO, R. *La psicologia sperimentale di Sante de Sanctis*. Roma: Orsa Maggiore, 1946.
2. BANISSONI, F. Psicologia sperimentale. In Various, *Un secolo di progresso scientifico italiano, 1839-1939*. Roma: S.T.P.S., 1939, Vol. IV. Pp. 377-427.
3. BANISSONI, F. Instituto Nazionale di Psicologia. *Ricerca Scient.*, 1950, **20**, 1625-1629.
4. BENUSSI, V. Die Psychologie in Italien. *Arch. gesam. Psychol.*, 1906, **7**, Literaturbericht, 141-180.
5. BOTAZZI, F. Agostino Gemelli et ses études psychologiques. *Scientia, Milano*, 1940, **67**, 115-124.
6. CHIABRA, G. The tendencies of experimental psychology in Italy. *Amer. J. Psychol.*, 1904, **15**, 515-525.
7. DE SANCTIS, SANTE. Autobiography. In C. Murchison (Ed.), *History of psychology in autobiography*. Vol. 3. Worcester: Clark Univer. Press, 1936. Pp. 83-120.
8. DYBOWSKI, M. *Dzialanie woli* (The activity of the will). Poznań: Ksiegarnia Akademicka, 1946.
9. FERRARI, G. C. Experimental psychology in Italy. *Amer. J. Psychol.*, 1905, **16**, 225-227.
10. FERRARI, G. C. Autobiography. In C. Murchison (Ed.), *History of psychology in autobiography*. Vol. 2. Worcester: Clark Univer. Press, 1932. Pp. 63-88.
11. FERNBERGER, S. W. Some European laboratories—1951. *Amer. J. Psychol.*, 1952, **65**, 619-626.
12. GEMELLI, A. Autobiography. In E. G. Boring, H. S. Langfeld, H. Werner, & R. M. Yerkes (Eds.), *History of psychology in autobiography*. Vol. 4. Worcester: Clark Univer. Press, 1952. Pp. 97-121.
13. GEMELLI, A., & BANISSONI, F. Speranze e preoccupazioni degli psicologi italiani in tema di insegnamento della psicologia nelle università italiane e nei vari tipi di scuole dell'ordine superiore. *Arch. Psicol. Neurol. Psychiat.*, 1941, **2**, 796-821.
14. IMUS, H. A. *Experimental psychology in Italy*. (Technical Report ONRL 63-52, Unclassified.) London: American Embassy, 1952.
15. KIESOW, F. Autobiography. In C. Murchison (Ed.), *History of psychology in autobiography*. Vol. 1. Worcester: Clark Univer. Press, 1930. Pp. 163-190.
16. LEMKAU, P. V., & DE SANCTIS, C. A survey of Italian psychiatry, 1949. *Amer. J. Psychiat.*, 1950, **107**, 401-408.
17. LORENZINI, G. L'Istituto di Psicologia Sperimentale del Pontificio Ateneo Salesiano. *Salesianum*, 1947, **9**, 240-258.
18. MANOIL, A. La psychologie expérimentale en Italie. *École de Milan*. Paris: Alcan, 1938.
19. MARZI, A. Scuola e psicologia. *Riv. Psicol.*, 1943, **39**, 166-172.
20. MARZI, A. La psicologia in Italia dal 1939 al 1943. *Riv. Psicol.*, 1944-45, **40-41**, 193-207.
21. MESCHIERI, L. *Bibliographia psicologica italiana: bollettino di informazione*, 1948, No. 1. Roma: Istituto Nazionale di Psicologia, 1951.
22. MURCHISON, C. (Ed.) *Psychological register*. Vol. 3. Worcester: Clark Univer. Press, 1932. Pp. 1053-1089.
23. NAVILLE, P. Quelques aspects du développement de la psychologie en Italie. *Enfance*, 1948, **1**, 449-453.
24. PLOTTKE, P. Psychology in Italy. *Indiv. Psychol. Bull.*, 1946, **5**, 89-91.
25. PONZO, M. In memoria di Alessandro Gatti. *Riv. Psicol. norm. pat.*, 1938, **34**, 1-2.
26. SAFFIOTTI, F. V. La evoluzione della psicologia sperimentale in Italia. *Riv. Psicol.*, 1920, **16**, 129-153.
27. WILGALIS, K. H. Die Stellung der Psychologie an Italienischen Universitäten und Instituten. *Z. Psychol.*, 1938, **142**, 193-199.
28. ZUNINI, G. *Psicologia*. (2nd Ed.) Brescia: Morcelliana, 1950. Chap. 15. La psicologia in Italia.

Received October 29, 1952.

## THE MEASUREMENT OF INDIVIDUAL DIFFERENCES IN ORIGINALITY<sup>1</sup>

R. C. WILSON, J. P. GUILFORD, AND P. R. CHRISTENSEN

*The University of Southern California*

One of the most important aspects of creative thinking is originality. This article discusses the problem of developing methods for measuring individual differences in originality. The problem arose in connection with a factor-analytic study of creative thinking conducted at the University of Southern California.<sup>2</sup>

In that investigation various definitions of originality were considered in the light of their implications for measurement. Three definitions and corresponding methods of measuring originality were finally adopted and applied to specially constructed tests. The methods are based upon: (a) uncommonness of responses as measured by weighting the responses of an individual according to the statistical infrequency of those responses in the group as a whole; (b) the production of remote, unusual, or unconventional associations in specially prepared association tests; and (c) cleverness of responses, as evaluated by ratings of degrees of cleverness exhibited in titles suggested for short-story plots.

These three methods permit the operations of measurement of individual differences and, while recasting the definition of originality, they preserve much of the essential meaning usually assigned to the concept. In the following sections, some of the

nonmeasurable aspects of originality are pointed out and each of the three proposed methods is discussed in conjunction with a description of tests developed to utilize the method. Since the tests were included in a factor analysis along with other tests of creative thinking, the three methods are evaluated in the light of the loadings of scores from these tests on a factor which has been called originality (2).

### DEFINITION OF ORIGINALITY

In developing methods for measuring individual differences in originality, the meaning to be assigned to the term *originality* and the operations for measurement must be clearly specified. The term originality has several distinct meanings. We wish to use it as the name for a psychological property, the ability to produce original ideas. What we mean by an original idea will be further specified in relation to each of the proposed methods of measuring originality.

Many writers define an original idea as a "new" idea; that is, an idea that "did not exist before." They are frequently not in agreement, however, in their interpretation of "new," since they use it with different connotations. We shall point out the inadequacy of two of these connotations for the measurement of individual differences in originality.

In one connotation, a "new" idea is an idea that "has never previously been thought of by anyone who has ever lived." In practice, of course, it would be impossible to verify whether

<sup>1</sup> Based in part on a paper presented at the Amer. Psychol. Ass., Chicago, August, 1951.

<sup>2</sup> Under Contract N6onr-23810 with the Office of Naval Research. The opinions expressed are our own and are not necessarily shared by the Office of Naval Research. (For a full account of this study see [1, 2].)

or not an idea meets these requirements of newness since one could never examine all the ideas of everyone who ever existed to determine whether the idea has been thought of before. This conception also presents a problem in the case of independent productions of the same idea. Two or more scientists may produce the same idea independently in different parts of the world. One of them may precede the others by a matter of months or weeks, or even hours or minutes. In trying to find creative scientists, we would probably not wish to regard the scientists who produced the idea later as unoriginal merely for having been preceded by someone unknown to them.

On the other hand, we find that "new," while meaning that which did not exist before, is sometimes interpreted, at least by implication, as including all human behavior that is not repetitive. That is, not only poetry, science, and inventions, but dreams, hallucinations, purposive behavior, and all perceptions are regarded as new. They are "new" in the sense that they are never duplicated exactly, even by the individual himself. Such a conception of "new" also fails to be fruitful, since it does not supply us with a basis for differentiating between more original and less original individuals.

For measurement purposes, we have found it useful to regard originality as a continuum. We have further assumed that everyone is original to some degree and that the amount of ability to produce original ideas characteristic of the individual may be inferred from his performance on tests. Rather than define original as "new" or "did not exist before" we have investigated three alternative definitions. We have regarded originality in turn as meaning "[uncom-

mon," "remote," and "clever." It was felt that these three definitions include significant aspects of what is commonly meant by the term original. Tests and scoring methods were developed for each of these approaches to originality.<sup>3</sup>

#### THE UNCOMMONNESS-OF-RESPONSE METHOD

Our first approach to the measurement of originality assumes a continuum of uncommonness of response. For this purpose originality is defined operationally as the ability to produce ideas that are statistically infrequent for the population of which the individual is a member. "Population" may here be regarded as any cultural group, professional group, or other aggregation of individuals having significant characteristics in common.

This definition of originality was utilized by constructing completion or open-end tests, which require the examinee to produce responses. The tests were administered to the group of individuals whose relative degrees of originality were to be determined. The responses of all the members of the group were tallied to determine their frequency of occurrence within the group. Weights were assigned to the various responses, the higher weights being given to the statistically more infrequent responses. A score was derived for each individual either by summing the weights assigned to his responses or by counting only the responses having high weights. On the basis of the score thus derived, those individuals with the highest scores were the individuals who had

<sup>3</sup> Other project personnel who contributed significantly, particularly to studies of the scoring procedures, are Raymond M. Berger, Norman W. Kettner, Donald J. Lewis, and Gordon Taaffe.

given the most infrequently mentioned responses.

This procedure may be clarified by an example. The items in the Unusual Uses test are six common objects. Each object has a common use, which is stated. The examinee is asked to list six other uses for which the object or parts of the object could serve. For example, given the item "A newspaper," and its common use, "for reading," one might think of the following other uses for a newspaper: (a) to start a fire, (b) to wrap garbage, (c) to swat flies, (d) stuffing to pack boxes, (e) to line drawers or shelves, (f) to make up a kidnap note. The test is given in two separately timed parts of five minutes each. Each part gives the names of three objects and their common use with spaces for listing six other uses per object.

All the responses given by a group of 410 Air Cadets and Student Officers to each object were classified, tallied, and weighted. A system of five weights was used. A weight of 5 was assigned for the (approximately)  $\frac{1}{2}$  most infrequently mentioned responses, a weight of 4 for the  $\frac{1}{2}$  next most infrequently mentioned responses, and so on down to a weight of 1 for the  $\frac{1}{2}$  most frequently mentioned responses. This gave a possible range of scores for each object (six responses) of 0 to 30 and a possible range of scores for the total test (six objects) of 0 to 180. The total scores actually obtained ranged from 5 to 129.

Let us consider the actual frequencies obtained for one of the objects. The 1,767 responses to the object given by the group of 410 Air Cadets and Student Officers were tabulated. One hundred and eighty-two different uses were mentioned. Eighty of these 182 uses were unique in that they were mentioned by only

one member of the group. At the other extreme, one of the uses was mentioned by 173 individuals. The three most common uses mentioned, with frequencies of 173, 94, and 90, accounted for 357 responses and were assigned weights of one. The next six most common uses, with frequencies from 89 to 48, were assigned weights of two. Nine uses with frequencies from 45 to 29 received weights of three, 24 uses with frequencies from 23 to 9 received weights of four, and the 139 most uncommon uses, with frequencies from 8 to 1, received weights of five. It should be noted that there were not exactly  $\frac{1}{2}$  of the total number of responses in each weight category. Because of the way in which the responses distribute themselves it is usually not possible to designate an exactly equal number of responses for each weight. It is possible, however, to achieve a close approximation.

After the weight for each response had been determined for all six objects, each examinee's paper was scored by assigning the appropriate weights to his responses and summing them. By definition, those individuals who tended to produce the most infrequently given ideas were the ones with the highest total scores and were regarded as the most original members of the group. The mean score on the Unusual Uses test was 64.0, its standard deviation was 23.5, and its alternate-forms reliability was .74.

The same procedure was applied to the Quick Responses test and the Figure Concepts test.<sup>4</sup> The Quick Responses test is similar to the conventional word-association test. It

<sup>4</sup> For more complete descriptions of tests mentioned here see (1).

consists of a list of 50 stimulus words, derived principally from the Kent-Rosanoff list and a more recent list developed by D. P. Wilson (4). The 50 words were read to the examinees at the rate of one every five seconds, the examinee being instructed to respond with the first word that came to mind. Responses of 410 individuals were tabulated for each of the 50 stimulus words. Frequencies of occurrence for each response were determined, weights were assigned, and scores derived in a manner similar to that for the Unusual Uses test. The mean score on the Quick Responses test was 99.8 with a standard deviation of 18.7. The reliability estimate was .81 as computed for odd and even items and corrected for length.

The Figure Concepts test consists of 20 simple pen-and-ink drawings of objects and individuals. Each picture is identified by a letter. The examinee's task is to find qualities or features that are suggested by two or more drawings and to list the features and the letter designations of two drawings which possess them. For example, picture A might be a sketch of a child wearing a hat, picture B might be a sketch of a woman wearing a hat, picture C might be a sketch of young birds in a nest. The examinee might give such responses as "wearing a hat (a, b)"; "young (a, c)"; "family (a, b)"; etc.

All responses of all individuals were tabulated and classified according to frequency of mention. A further breakdown was made for each response mentioned in terms of the combinations of drawings used in identifying the feature. It was noted that while there were 190 possible pairs of drawings available, certain ones were rarely used, while others were used as a source of more than

one feature. Weighting of responses was thus based on both the infrequency of the response itself and the infrequency of the drawing combination used as a source of that response.

How this dual classification affected an individual's score may be seen in the situation where two individuals gave the same response (feature name), but cited different combinations of drawings. If one individual's response was derived from a drawing combination that was frequently mentioned by others in connection with that feature, the weight assigned was low. The other individual's response, if derived from a drawing combination infrequently mentioned for that feature, was assigned a high weight.

As with the Unusual Uses test, weights were assigned so that an approximately equal number of all the responses given by the group received each weight. Each examinee's responses were then assigned their appropriate weights and the weights were summed to derive the individual's total score for the test. The mean score on this test was 29.9 with a standard deviation of 12.9. Since the format of this test did not permit the direct computation of a reliability estimate, the communality of the test (.41) found in the factor analysis is offered as an estimate of a lower bound of its reliability.

In the Number Associations test the examinee is given, in turn, four different numbers (digits) and for each is allowed two minutes in which to list as many synonyms, uses, and things associated with the number as he can. For example, for the number 4 he might list coach-and-four, for, fore, foursome, quartet, etc.

The associations listed by the group were tabulated and weights were assigned in a manner similar to

that described for the Unusual Uses test. In order to try out a further variation of the uncommonness method, however, the individual's total score was derived in a slightly different manner from that previously described. Instead of summing the weights for all the responses given by the individual, his total score was derived by counting the number of responses with weights of 4 and 5. The mean score for this test was 12.5 with a standard deviation of 3.6 and an alternate-forms reliability of .57.

In the approach described in this section, we have chosen to define original as meaning "uncommon." An original idea or response is one that is uncommon or statistically infrequent, and an individual's degree of originality, as inferred from his scores on the tests described, is characterized by the degree of uncommonness of his responses.<sup>6</sup>

#### THE REMOTENESS-OF-ASSOCIATION METHOD

The second approach is in terms of remoteness of association. Originality is here defined as the ability to make remote or indirect associations. To measure originality from this point of view, tests were constructed that required the examinee to make remote associations if he responded at all. Remoteness of association was imposed by the task. Three tests of this type were constructed. The degree of originality of an individual, ac-

<sup>6</sup> The reader may recall that an uncommonness or idiosyncrasy score has previously been used in connection with word-association tests in the assessment of abnormalities of behavior in clinical practice, particularly of the schizoid type. The fact that such a score measures an originality factor, as we shall show later, might be regarded as support for the popular idea expressed in the words of Seneca, "There is no great genius without some touch of madness."

cording to this definition, would be manifested in terms of the number of remote associations he made.

The Associations I test presents 25 pairs of words. The associative connection between the two words is not immediately apparent. The examinee's task in each item is to call up a third word that serves as a link between them. For example:

Given:

Indian \_\_\_\_\_ money

Write on the line between these words a word that associates the two.

There are several possible words that could be used such as penny, nickel, copper, and wampum, each of which is related to both Indian and money.

The examinee's score was the number of responses given to the 25 items in four minutes. The mean score for this test was 14.0 with a standard deviation of 4.9. The odd-even reliability estimate was .87, corrected for length.

The Associations II test is similar to the Associations I except that there is more emphasis on the correct response word having two different meanings in its relationship to the two stimulus words. It is also a multiple-choice test in which the examinee must indicate which one of five letters is the first letter of the correct association.

For example:

tree a b g m s dog

Which of the five letters is the first letter of a word that is associated with both tree and dog and has a different meaning in relation to each?

The word "bark" is the correct answer. It means the external covering of a tree and it also means the noise made by a dog. It also begins with *b* which is one of the choices, so the examinee circles the letter *b*.

The examinee's score was the number of correct responses given to 25 items in 12 minutes. The mean score was 14.0 with a standard deviation of 3.9. The odd-even reliability estimate was .62, corrected for length.

The Unusual Uses test, previously described, was also regarded as a test requiring the examinee to respond with remote associations. Since the six items composing the test were common objects, each with one well-known use, which was given, the examinee was compelled to utilize remote associations in seeking six additional uses for each object. Both a statistical-infrequency score and a simple-enumeration score were derived for this test. The correlation between these two scores was .94. There is, of course, much spurious overlap of the two scores. In view of the high correlation between the two scores and the similarity of their correlations with other tests in the creative-thinking battery, the simpler score was chosen for inclusion in the factor analysis. The mean for this score on this test was 22.1 with a standard deviation of 6.7 and an alternate-forms reliability of .80.

In the approach described in this section we have chosen to define original as meaning "remote." An original idea or response is "remote" to the extent that the individual is required to bridge an unusually wide gap in making associative responses. An individual's relative originality, as inferred from his scores on these tests, is characterized by the number of remote responses given in limited time.

#### CLEVERNESS

According to the third approach, originality is defined as the ability to produce responses that are rated as clever by judges. This definition

requires a test that calls forth responses showing variation on a continuum of cleverness. Weights are assigned to an individual's responses in proportion to their degrees of rated cleverness.

The Plot Titles test used to measure this type of originality presents two brief stories. For each story the examinee is allowed three minutes in which to write as many appropriate titles as he can. Although relevancy rather than cleverness is stressed in the instructions, an examination of the responses of the group revealed considerable variation in the ingenuity, cleverness, or striking quality of the titles suggested.

In an attempt to develop a reliable scoring procedure for evaluating cleverness, a sample of 50 individuals was selected from the total group of 410. These 50 individuals averaged approximately six responses for each plot. The approximately 300 titles for each plot were typed on separate slips of paper. Three judges, working independently, sorted the titles into six successive piles on the basis of their judgments of the relative cleverness of the titles. Weights from 0 through 5 were assigned to the titles in the successive piles, with the high weights being assigned to the more clever titles. Agreement among the judges is indicated by the interjudge correlations (of ratings) ranging from .53 to .76. Reliabilities of test scores derived from individual judges ranged from .69 to .77. These reliabilities were computed from the two cleverness scores, one from each story, for each of the 50 individuals. The reliability computed from the composite ratings of the three judges (.76) was not higher than that for the best individual judge. Since the most reliable judge was also the one who agreed best with the other two

judges, it was decided to have this one judge do the scoring of the test for all examinees, with one of the other judges serving as a check scorer.

In an effort to simplify scoring, a study was made of total scores derived from the weights 0 through 5. That is, each test paper was scored by the number of responses at each of the cleverness levels of 0, 1, 2, 3, 4, and 5. Intercorrelations among the six scores were computed for the sample of 50 individuals. It was found that scores based on weights 0 and 1 intercorrelated well, and scores based on weights 2, 3, 4, and 5 intercorrelated well. A combination of scores based on weights 0 and 1 had a low correlation with a combination of scores based on weights 2, 3, 4, and 5. It was decided to reduce the scale to two intervals, clever and nonclever. That is, responses receiving weights 0 and 1 would be called nonclever. This greatly reduced the fineness of discrimination required of the scorer. Utilizing the titles already rated as a standard, the remainder of the tests were scored on this simple dichotomy. Two scores were recorded for each individual: the number of clever titles and the number of nonclever titles. It was decided to include both scores in the computation of the intercorrelation matrix and to determine, prior to the factor analysis, whether the cleverness and noncleverness scores were sufficiently independent to warrant including both of them in the factor analysis. The correlation between the two scores was  $-.031$  and their patterns of intercorrelations with other tests in the battery were quite different; consequently, both scores were included in the factor analysis. The cleverness score (based on weights 2 to 5) emerged with a loading of .55 on the orig-

inality factor. The noncleverness score (weights 0 to 1) had a loading of  $-.05$  on this factor and had its highest loading (.59) on a factor identified as ideational fluency. The cleverness score had a loading of .07 on the ideational-fluency factor.

In the approach described in this section, we have chosen to define original as meaning "clever." An original idea or response is one that is rated as clever by judges. An individual's degree of originality, as inferred from this kind of test score, would be characterized by the number of clever responses given in limited time.

#### DISCUSSION

The seven test scores representing the three scoring methods described were included with 46 other test scores in a battery designed to explore the domain of creative thinking. The test battery was administered to 410 Air Cadets and Student Officers. The scores were intercorrelated and 16 factors were extracted. Orthogonal rotations resulted in 14 readily identifiable factors, a doublet, and a residual. Five of the seven originality test scores emerged with loadings regarded as significant (.30 and above)<sup>6</sup> on one of the factors obtained. Following is a list of the tests, their scoring principles, and their loadings on the factor.

Plot Titles (cleverness)	.55
Quick Responses (uncommonness)	.49
Figure Concepts (uncommonness)	.32
Unusual Uses (remoteness)	.31
Associations I (remoteness)	.30
Number Associations (uncommonness)	.25
Associations II (remoteness)	.09

<sup>6</sup> The common practice among factor analysts is to regard only loadings greater than .25 or greater than .30 as significant. That is, since no means are available for determining

We have tentatively named this factor originality.<sup>7</sup> Another test from the creative thinking battery which should be discussed in relation to this factor is the Consequences test. This test requires the examinee to list the consequences of certain unexpected events such as the sudden abolition of all national and local laws. Two scores were derived from this test on the basis of the degree of remoteness of ideas indicated by the individual's responses. The number of remote consequences was counted for one score and the number of immediate or direct consequences for the other. It was hypothesized that the remoteness of ideas represented by the remote-consequences score might refer to something different from the remoteness of ideas required by the originality tests already mentioned. A separate factor of penetration or the ability to see remote consequences in space, in time, or in a causal chain of circumstances was therefore hypothesized. No such factor emerged in the factor analysis. The remote-consequences score of the Consequences test came out with its highest loading (.42) on the originality factor. Evidently, the remoteness of ideas represented by this test score is not different from the remoteness of ideas required by the test scores hypothesized for originality. This finding lends additional

support to the generality of the obtained originality factor.

Inasmuch as test scores representing all three methods of measuring originality have significant loadings on this factor, we may have some confidence in its generality. Had test scores of only one method emerged on the factor, we might wonder whether the factor were specific to the particular kind of scoring method.

It should be mentioned that this factor has some appearance of bipolarity since there were a few small negative loadings of other test scores in the battery on this factor. Those test scores with negative loadings are of the kind whose "right" responses are keyed on an arbitrary, conventional basis by the test constructor. The examinee who engages in an unusual line of thought is likely to be penalized for his originality in such tests. In this connection, the essentially zero loading for originality in Associations II (as contrasted with the significant loading in Associations I) is worth mentioning. In this test, too, one "correct" answer is given credit. It may be that the original examinees think of other appropriate responses whose initial letters appear among the alternatives, and for which they receive no credit.

The fact that five of our tests designed to measure originality have in common a single factor is regarded as evidence for the potential fruitfulness of the scoring methods described for the measurement of individual differences in originality. Further work is necessary in refining the tests and in validating them against objective criteria of originality. It is felt that considerable progress has been made toward the development of objectively scored tests of origi-

the standard error of a factor loading, it has arbitrarily been decided to interpret the nature of factors on the basis of those loadings greater than the indicated cutoffs. The more conservative one (.30) is used by the present investigators.

<sup>7</sup> Hargreaves (3), in a study of imagination tests using Spearman's method of tetrad differences, found an originality group factor based on tests scored for uncommonness in a fashion similar to that described earlier in this report.

nality, with promise of satisfactory reliability.

As to the relative merits of the three approaches suggested, the uncommonness and cleverness methods have the greatest amount of the originality-factor variance but are the least economical in time and energy required to determine the scores.

In an exploratory study such as this one, expenditure of time and energy in scoring by the less economical methods may be justified in terms of the insights to be gained.

In later studies, however, it is desirable to use more economical procedures. The remoteness principle is a more economical procedure, but does not yield factor loadings as high as the less economical cleverness and uncommonness procedures. The next steps will be to revise the remoteness tests in an attempt to increase their originality variance and to seek methods of simplifying further the cleverness and uncommonness scoring procedures without decreasing their originality variance.

#### REFERENCES

1. GUILFORD, J. P., WILSON, R. C., CHRISTENSEN, P. R., & LEWIS, D. J. A factor-analytic study of creative thinking, I. Hypotheses and description of tests. *Reports from the Psychological Laboratory*, No. 4. Los Angeles: Univer. of Southern California, 1951.
2. GUILFORD, J. P., WILSON, R. C., & CHRISTENSEN, P. R. A factor-analytic study of creative thinking, II. Administration of tests and analysis of results. *Reports from the Psychological Laboratory*, No. 8. Los Angeles: Univer. of Southern California, 1952.
3. HARGREAVES, H. L. The "faculty" of imagination. *Brit. J. psychol. Monogr. Suppl.*, 1927, 3, No. 10.
4. WILSON, D. P. An extension and evaluation of association word lists. Unpublished doctor's dissertation, Univer. of Southern California, 1942.

Received January 19, 1953.

## CORRECTING THE KUDER-RICHARDSON RELIABILITY FOR DISPERSION OF ITEM DIFFICULTIES

PAUL HORST

*University of Washington*

Loevninger (5) has insisted, quite rightly, that the Kuder-Richardson (4) reliability formulas are actually estimates of item homogeneity as well as of test reliability. She proceeds, then, to point out that K-R 20<sup>1</sup> has unity as an upper limit only when the items in the test are all of equal difficulty, and regards as a serious defect of formula K-R 20 the fact that its upper limit is a function of the dispersion of item difficulties. She argues that a test of perfectly homogeneous items is justified only if there is a range of difficulty in the items. Otherwise, a single item will give the same discrimination as any number of equally difficult perfectly homogeneous items. Therefore, she contends, what is needed is a coefficient of homogeneity which has unity as its maximum value, irrespective of the dispersion of item difficulties. Accordingly, she develops a formula which has this property. It is given in somewhat different notation by

$$H_t = \frac{\sigma_t^2 - \Sigma pq}{\sigma_m^2 - \Sigma pq} \quad [1]$$

where  $H_t$  is the coefficient of homogeneity of the test;  $\sigma_t^2$  the test variance;  $\sigma_m^2$  is the maximum variance possible for a test which has the same distribution of item difficulties as the test under consideration;  $\Sigma pq$  is the sum of the item variances in the test. It is clear, therefore, that as  $\sigma_t^2$  ap-

proaches  $\sigma_m^2$ ,  $H_t$  approaches unity. It has been shown by Carroll (1) in somewhat different notation that the maximum possible value of the variance of a work limit test, with score being number of items correct, is given by

$$\sigma_m^2 = 2\Sigma ip_i - M_t(1 + M_t) \quad [2]$$

where  $M_t$  is the test mean and  $\Sigma ip_i$  means that the  $p$ 's are ranked in descending order of magnitude, each  $p$  is multiplied by its rank order, and the sum of the products is taken.

Suppose now we attempt to relate equation [1] to one of the more familiar reliability functions. Kuder and Richardson have shown (4) that the reliability of a test is given by

$$r_{tt} = \frac{\sigma_t^2 - \Sigma pq + \Sigma r_{ii}pq}{\sigma_t^2} \quad [3]$$

where  $r_{tt}$  is the reliability of the test;  $\sigma_t^2$  is the variance of the test;  $p$  is the difficulty of an item;  $q = 1 - p$ ;  $r_{ii}$  is the reliability of item  $i$ .

The problem in using [3] is to get plausible estimates of the item reliabilities. Kuder and Richardson in deriving their well-known formula 20 assumed, in effect, that all item reliabilities are equal. Although they proceeded somewhat differently, we may, from this assumption, rewrite [3] as

$$r_{tt} = \frac{\sigma_t^2 - (1 - r_{ii})\Sigma pq}{\sigma_t^2} \quad [4]$$

If now we let  $r_{ii}$  be the ratio of the average item covariance to the aver-

<sup>1</sup> This formula is

$$r_{tt} = \frac{n}{n-1} \left( \frac{\sigma_t^2 - \Sigma pq}{\sigma_t^2} \right).$$

age item variance, we have shown elsewhere (2) that

$$r_{ii} = \frac{\sigma_i^2 - \Sigma pq}{(n-1)\Sigma pq}. \quad [5]$$

Except for a slight difference in notation, equation [5] gives the same estimate of item reliability which Kuder and Richardson derived by somewhat different methods in their formula 18. Substituting [5] in [4], we get

$$r_{ii} = \frac{n - \left[ 1 - \left( \frac{\sigma_i^2 - \Sigma pq}{(n-1)\Sigma pq} \right) \right] \Sigma pq}{\sigma_i^2} \quad [6]$$

which reduces to

$$r_{ii} = \frac{n}{n-1} \left( \frac{\sigma_i^2 - \Sigma pq}{\sigma_i^2} \right). \quad [7]$$

Equation [7] is, of course, the well-known Kuder-Richardson formula 20 for estimating test reliability.

Let us now examine equation [5] in the light of Loevinger's criticism. For a set of items of specified difficulties,  $r_{ii}$  will be a maximum only when  $\sigma_i^2$ , or the variance, is as large as possible. This maximum value,  $\sigma_i^2$ , is given by equation [2]. First we shall write [5] in the form

$$r_{ii} = \frac{\sigma_i^2 - \Sigma pq}{n\Sigma pq - \Sigma pq}. \quad [8]$$

We shall prove that [8] cannot be unity unless all of the items are of equal difficulty. It is well known that the variance of a test can be expressed as a function of the item variances and the interitem correlations, if the score is number correct and the test is not speeded. This function is

$$\sigma^2 = \left( \begin{array}{c} \sigma_1^2 + s_1 s_2 r_{12} + \dots + s_1 s_n r_{1n} \\ + s_1 s_2 r_{12} + \sigma_2^2 + \dots + s_2 s_n r_{2n} \\ \dots \dots \dots \dots \dots \dots \\ + s_n s_n r_{1n} + s_2 s_n r_{2n} + \dots + s_n^2 \end{array} \right) \quad [9]$$

where  $S_i = \sqrt{p_i q_i}$  and the  $r_{ij}$  are phi coefficients. If we prove that

$$\sigma_i^2 < n\Sigma pq \quad [10]$$

for items of unequal difficulty, we shall have proved that  $r_{ii}$  is less than unity, since the last terms of both numerator and denominator of [8] are equal. Since  $S_i = \sqrt{p_i q_i}$  we must prove that

$$\sigma_i^2 < n\Sigma s_i^2. \quad [11]$$

It can readily be shown that a phi coefficient must be less than unity unless both items are equally difficult. In no case can they exceed unity. If we prove that, even though all  $r$ 's in [9] are unity, the inequality in [11] must still hold for unequal  $S$ 's, then we will also have proved it for the case of some  $r$ 's less than unity. Assuming all  $r$ 's in [9] unity, we write

$$\sigma_i^2 = (\Sigma s)^2. \quad [12]$$

Substituting [12] in [11],

$$(\Sigma s)^2 < n\Sigma s^2. \quad [13]$$

We indicate the variance of the  $s$ 's by  $\sigma_s^2$  and, from the standard formula, write

$$\sigma_s^2 = \frac{\Sigma s^2}{n} - \left( \frac{\Sigma s}{n} \right)^2$$

or

$$(\Sigma s)^2 = n\Sigma s^2 - n^2\sigma_s^2. \quad [14]$$

Substituting [14] in [13],

$$n\Sigma s^2 - n^2 - n^2\sigma_s^2 < n\Sigma s^2 \quad [15]$$

or

$$n^2\sigma_s^2 > 0. \quad [16]$$

Obviously, [16] holds unless the variance of the item sigmas is zero. But the variance of the item sigmas cannot be zero unless all the item difficulties are equal,

or unless some of them are equal to a constant  $p$  and all the others are equal to a constant  $1-p$ . But in the latter case not all of the phi coefficients could be unity. So the only case in which  $\sigma_i^2$  could be as large as  $n \Sigma pq$  is when the item variances are all equal and the phi coefficients are all unity, and this can be true only when all items are of equal difficulty.

Loevinger (6) and Johnson (3) have proposed that, in order to make a phi coefficient independent of the disparity of its item difficulties, we should divide it by the maximum phi which is possible for the two obtained difficulties. Generalizing from this rationale, we shall apply the procedure to  $r_{ii}$  in [5].

$$r_m = \frac{\sigma_m^2 - \Sigma pq}{(n-1)\Sigma pq} \quad [17]$$

where  $\sigma_m^2$  is the maximum variance of a test with specified item difficulties. We then let  $r_{ii}$  be the adjusted value of  $r_{ii}$  and write

$$r_{ii} = \frac{r_{ii}}{r_m} \quad [18]$$

or substituting [5] and [17] in [18],

$$r_{ii} = \frac{\sigma_i^2 - \Sigma pq}{\sigma_m^2 - \Sigma pq} \quad [19]$$

But we see that equation [19] is precisely equation [1], or Loevinger's coefficient of homogeneity. This is the coefficient which she preferred to K-R 20. However, this preference does not seem entirely appropriate. As we have seen, it gives an estimate of average item intercorrelation corrected for dispersion of item difficulties. Loevinger's proposal lacks one essential step. To get a more realistic estimate of test reliability, we should substitute her formula for the estimate of average item reliability in equation [4]. Substituting equation

[19] in [4] we get

$$r_{ii} = \frac{\sigma_i^2 - \left(1 - \frac{\sigma_i^2 - \Sigma pq}{\sigma_m^2 - \Sigma pq}\right) \Sigma pq}{\sigma_i^2} \quad [20]$$

which reduces to

$$r_{ii} = \left( \frac{\sigma_i^2 - \Sigma pq}{\sigma_m^2 - \Sigma pq} \right) \frac{\sigma_m^2}{\sigma_i^2} \quad [21]$$

Equation [21] would seem to provide a more realistic estimate of the reliability coefficient than K-R 20. It yields a higher estimate if the items are of unequal difficulty and will have unity as an upper limit irrespective of the dispersion of item difficulties. Because of the factor  $\sigma_m^2/\sigma_i^2$ , it yields a higher value than Loevinger's homogeneity coefficient given by [1].

The chief disadvantage of using [21] rather than K-R 20 is, of course, the added time required to compute  $\sigma_m^2$  as given by [2]. However, the major part of the labor involved in using K-R 20 is getting the item counts for the  $pq$  values. Once the  $p$ 's are obtained, it is not a great deal more work to get  $\Sigma ip$  required in equation [2] to get  $\sigma_m^2$ .

If we have a wide distribution of item difficulties, the difference between K-R 20 and our equation [21] may actually be quite large. Suppose, for example, we have for seven items in ascending order of difficulty, the following values: .80, .70, .60, .50, .40, .30, and .20. Suppose the variance of the test is 5.25. We have, then, the following values:

$$n = 7$$

$$M_t = \Sigma p = 3.50$$

$$\sigma_i^2 = 5.25$$

$$\Sigma pq = 1.470$$

$$\Sigma ip = 11.20$$

For the K-R 20 value we have

$$r_u = \frac{7}{6} \left( \frac{5.25 - 1.47}{5.25} \right) = .84.$$

For the adjusted value given by our equation [21] we must first compute the maximum variance given by equation [2]. This is

$$\sigma_m^2 = 2 \times 11.20 - (3.50)(4.50) = 6.65.$$

Substituting the required values in [21], we have

$$r_u = \left( \frac{5.25 - 1.47}{6.65 - 1.47} \right) \frac{6.65}{5.25} = .92.$$

In this particular case, therefore, with a rectangular distribution of item difficulties ranging from .80 down to .20, our estimate of reliability is almost 10 per cent higher than that given by K-R 20. It should be emphasized that, even with the correction we suggest, we are compensating only for the attenuation introduced by the dispersion of item diffi-

culties. We still have to remember that the basic assumption in all these formulas is that all items measure the same function, and that the failure of maximum item intercorrelation is due only to the unreliability of the items. Thus, even our formula [21] should be regarded as only a lower bound to the Kuder-Richardson type of reliability even though this lower bound may be somewhat higher than that given by K-R 20.

It should be emphasized that the type of reliability which we refer to is consistency of behavior within a very limited time interval, that is, the time interval during which the items in the test are being responded to. This is, of course, also the type of reliability implied by the Kuder-Richardson formulas. However, it should also be observed that this "short term" reliability is comparable to the correlation between two tests taken within a short time interval, e.g., at the same sitting or on the same day.

#### REFERENCES

1. CARROLL, J. B. The effect of difficulty and chance success on correlation between items or between tests. *Psychometrika*, 1945, **10**, 1-19.
2. HORST, P. Relationships between several Kuder-Richardson reliability formulas. *J. educ. psychol. Measmt.*, in press.
3. JOHNSON, H. M. Maximal selectivity, correctivity and correlation obtainable in  $2 \times 2$  contingency-tables. *Amer. J. Psychol.*, 1945, **58**, 65-68.
4. KUDER, G. F., & RICHARDSON, M. W. The theory of the estimation of test reliability. *Psychometrika*, 1937, **2**, 151-160.
5. LOEVINGER, JANE. A systematic approach to the construction and evaluation of tests of ability. *Psychol. Monogr.*, 1947, **61**, No. 4 (Whole No. 285).
6. LOEVINGER, JANE. The technique of homogeneous tests. *Psychol. Bull.*, 1948, **45**, 507-529.

Received July 28, 1952.

## MODELS FOR TESTING THE SIGNIFICANCE OF COMBINED RESULTS<sup>1</sup>

LYLE V. JONES AND DONALD W. FISKE

*University of Chicago*

Experimenters in psychology frequently encounter the problem of assessing the significance of a set of experimental results, each of which has been tested for statistical significance. Although several recent articles (1, 2, 13) discuss the problem, none enunciates fully the appropriate models. Nor is the problem adequately covered in any commonly used statistical text. It is the aim of the present paper to clarify the differences between the models and to provide explicit statements of the assumptions underlying them and the situations to which they apply.

The general problem can be stated simply: an experimenter has a set of two or more experimental results. For each he has applied an appropriate test of significance. He then wishes to apply a test of significance to the entire set: he wishes to formulate and test a single hypothesis about the entire set of results. One example is the evaluation of several findings, each of which is relevant to the same general hypothesis. The findings may be viewed as repetitions of essentially the same experiment. Another example is the evaluation of a set of results from the same sample, such as the analysis of test items against a criterion or the validation of several predictors against one or more criteria. Both examples illustrate types of problems which arise frequently in

psychological research, where repetition of experiments and multivariate prediction are not uncommon.

The two applicable statistical models are outlined below.

### THE MODELS

*The binomial model.* Application of this model provides an evaluation of the significance of a set of results, one or more of which have been found to be significant. To illustrate, an experimenter has five results, two of which are significant at his pre-selected significance level (typically .05 or .01). He now wishes to test the hypothesis that the two significant results out of five could be expected by chance, against the alternative hypothesis that the proportion of significant results exceeds that expected on a chance basis.

The probability of one or more significant results can be determined from the binomial distribution, the expansion of  $(p+q)^N$  where  $p$  is the specified level of significance,  $q=1-p$ , and  $N$  is the total number of tests of significance. From the binomial, we find the probability of obtaining by chance  $n$  or more "successes" (results beyond the specified value of  $p$ ). The value of this probability is the sum of  $N-n+1$  terms of the binomial expansion:

$$\sum_{s=n}^N \binom{N}{s} p^s q^{N-s}.$$

Wilkinson (13) provides tables of this probability for  $p=.05$  and  $p=.01$ , for values of  $N$  up to 25. For

<sup>1</sup> The authors are indebted to a number of colleagues at the University of Chicago, and especially to Dr. John M. Butler and to Dr. Julius Seeman, for their constructive criticisms of an early form of this paper.

larger values of  $N$ , up to 50, and for other values of  $p$ , the *Tables of the Binomial Probability Distribution* (15) can be consulted in order to escape unwieldy computations. As is suggested by Brozek and Tiege (2), the normal approximation to the binomial distribution supplies a satisfactory solution so long as the product of  $N$  and  $p$  exceeds 5. Following this rule, the approximation seriously departs from the exact binomial solution for any  $N$  (number of independent statistics or results) smaller than 100 if one is using the .05 level of significance, or smaller than 500 for the .01 level.

The fundamental assumption of the binomial model is that the several experimental results are independent, that the probability value for any one result in no way influences the value for any other result. This assumption will be examined in detail below.

*The chi-square model.* This model is used to test the hypothesis that the composite  $p$  value for the several findings could have occurred by chance. It utilizes the numerical  $p$  values instead of classifying them as above or below a specified critical level of significance. This model also assumes the statistical independence

The model is based on the proof that any  $p$  value can be transformed to a chi-square value with two degrees of freedom, and that the sum of independent chi squares is distributed as chi square. The transformation equation is

$$x^2 = -2 \log_e p$$

or

$$x^2 = -2 (2.3026) \log_{10} p.$$

The composite  $x^2$  is given by the formula

$$x^2 = -2 \sum_{i=1}^k \log_e p_i$$

with  $2k$  degrees of freedom, where  $k$  is the number of independent probability values to be combined. Since

$$\sum_{i=1}^k \log_e p_i = \log_e (p_1 p_2 p_3 \cdots p_k),$$

the formula for the composite  $x^2$  may also be written

$$x^2 = -2 \log_e (p_1 p_2 p_3 \cdots p_k).$$

The product of the  $k$  separate  $p$  values is the joint probability of the  $k$  independent findings. There is no convenient test for directly assessing the significance of a joint probability.

TABLE 1  
EXAMPLE OF THE CHI-SQUARE TRANSFORMATION

$p$	$\log_e p$	$\log_{10} p$
.04	-3.2189	8.6021-10.0000
.05	-2.9958	8.6990-10.0000
.20	-1.6094	9.3010-10.0000
	<u>-7.8241</u>	<u>26.6021-30.0000</u>
$x^2 = -2 (-7.8241) = 15.6482$		
$x^2 = -2(2.3026) (-3.3979)$		
$x^2 = 15.6480$		
$df = 2 (3) = 6$		
$p < .02$		

of the several results being combined.

Suppose we have obtained the following set of  $p$  values from three com-

parable experiments: .04, .05, and .20. The composite  $\chi^2$  most conveniently can be computed using the summation formula above or the corresponding formula for common logarithms as exemplified in Table 1.

The proof for the adoption of this composite test is derived and developed by Karl Pearson (12) and by E. S. Pearson (11). It is illustrated by Kendall (8, pp. 132-133) and by Lindquist (10, pp. 46-47). Baker (1) provides an abac for determining a combined probability from the probabilities of two separate results. Gordon, Loveland, and Cureton (5) have recently published a table of  $-2 \log p$  for values of  $p$  from .001 to .999 which facilitates the application of this test.

An essential difference between the two models appears when we consider the amount of information utilized. In the application of the binomial model, use is made of a simple alternative: each result in the set either is significant or it is not. The application of the chi-square model, on the other hand, demands knowledge of the exact probability value associated with each result.

An experimenter may prefer to apply the binomial method to situations where the chi-square model is appropriate. He might wish to test a hypothesis on the basis of combined results by determining whether from among  $N$  independent tests of hypothesis,  $n$  or more significant results would be expected to appear by chance. However, this procedure would involve considerable risk of erroneous interpretation, when compared with the application of the chi-square model. For example, if he were operating at the .05 level of confidence and if the  $p$  values of three results were .09, .06, and .04, he would accept the null hypothesis,

for the probability of obtaining at least one result out of three beyond the .05 level is .14. But using the chi-square transformation, the combined  $p$  is .02. On another occasion, again working at the .05 level of confidence, three results might be associated with  $p$  values of .05, .04, and .95. Under the binomial model, the null hypothesis would be rejected, for the probability of achieving at least two significant results out of three is much smaller than .05. (Wilkinson's table shows it to be .0072.) The chi-square transformation, on the other hand, leads to a combined  $p$  slightly greater than .05. Such examples indicate the major differences between the two procedures and show that the binomial method, by asking a different question, ignores some of the information contained in the set of obtained  $p$  values.

#### ASSUMPTIONS AND LIMITATIONS

The most limiting assumption underlying interpretation of results from the application of either model, and the assumption most likely to be unwarranted, is that of statistical independence. Independence is a necessary condition in order that, under the null hypothesis, the  $N$  results be distributed as the binomial distribution in the first model and as chi square in the second. Independence, in this sense, means that the probability associated with each one of the  $N$  results is unrelated to the probability associated with each of the other  $N-1$  results. The finding that one of the results is significant must not affect the expected probability for any of the other results, which remains at .50, under the null hypothesis.

The effect of the assumption of statistical independence of data upon applications of the models can best

be assessed if we consider two kinds of situations to which the models might be applied.

Situation A. Each of the several experimental results is obtained from a *different* sample of individuals. Typically, in this situation, the several tests of a hypothesis represent replications based upon independent samples. To the extent that the samples *are* independent, either model may be applied. For the case where the samples have been randomly selected, their independence is assured.

Situation B. Each of several experimental results is obtained from the *same* sample of individuals. A typical case is item analysis, where interest resides in the degree of relationship between each of a number of items and a single criterion. The items are administered to a group of respondents, and the correlation between the criterion variable and each item is determined. Under the binomial model, the experimenter then asks if the number of significant item-criterion correlation coefficients exceeds the number expected to be significant by chance. The chi-square model would seldom be useful in this situation; when used, it would resolve the question whether the several findings support the contention that in this battery the distribution of item-criterion correlations is different from that expected by chance.

It is in dealing with this situation that our inquiry is often not legitimate, for it is seldom that the separate results are statistically independent. In the item-analysis example, independence presumes that the interitem correlations are randomly distributed around an expected value of zero, i.e., the items must be uncorrelated. Referring once again to the paper by Brozek and Tiede, we

are compelled to doubt the legitimacy of using the binomial model (or, indeed, either model) for the item-analysis problem which they illustrate. No evidence is provided which would indicate that the 228 questionnaire items analyzed in their report are independent. It seems unlikely that the thousands of interitem relationships were evaluated. Since the items had been selected to focus "on a variety of personality aspects described in the literature as being characteristic of individuals who are prone to develop high blood pressure" (2, p. 339), nonzero interitem correlations surely would be expected. Not only is it a serious methodological error to apply the binomial model in this example, but also the actual conclusion is likely to be wrong. It is impossible to accept the authors' conclusion that there is "a negligible probability" of obtaining 24 or more statistics, significant beyond the .05 level, in a series of 228 item statistics. With nonzero interitem correlations, the binomial model, as applied by Brozek and Tiede, is an invalid test.

Wilkinson talks of the experimenter who tests all possible relationships in his data, possibly correlating every variable with every other one. While he apparently condones the use of the binomial method and his tables for this situation, such an application also violates the assumption. The effect of the violation, however, is considerably less extreme than in the item-analysis example. Consider an array of  $n(n-1)/2$  intercorrelation coefficients among  $n$  variables. Under the conditions of the null hypothesis, these may be considered randomly selected from a sampling distribution with a population correlation of zero. The lack of independence of sample statistics may be seen when two or more particular

sample  $r$ 's depart from zero: e.g., if  $r_{ai}$  and  $r_{aj}$  each are positive, the expected value of  $r_{ij}$  is no longer zero, but becomes a small positive number. If the null hypothesis obtains, the effects of dependence are slight. The binomial test, if interpreted with caution, may serve as an approximate index of independence among the  $n$  variables.

This contention is supported by the results of three sampling experiments conducted by one of the writers. Scores on each of 20 artificial variables were drawn from tables of normalized random numbers (3, pp. 295-304 and 350-359). The number of scores per variable was 50, 100, and 200, respectively, for the three experiments. In each study IBM methods were used for computation of the intervariable correlation matrix. In every case the distribution of the 190 sample  $r$ 's was compared with the distribution expected under the hypothesis that these  $r$ 's had been independently and randomly selected from a sampling distribution with a population correlation of zero. Several tests of this hypothesis yield no evidence against it; e.g., the proportion of  $r$ 's exceeding each of several levels of significance did not significantly depart from that expected by chance. The proportion of positive  $r$ 's, as contrasted with the number of negative values, suggests no extrachance effects. In short, results from none of the three studies suggest that the nonindependence of obtained correlation coefficients was a factor which noticeably disturbed the sampling distribution.

In general, when combining results by either the binomial model or the chi-square transformation, the assumption of statistical independence is one which cannot be treated lightly. Some assumptions, such as that of

normally distributed variates underlying Student's  $t$  and Fisher's  $F$ , have been shown to be of little practical importance; those tests are relatively insensitive to moderate departures from normality. In contrast, if statistical dependence is present when results are being combined, the findings from either the binomial test or the chi-square transformation are likely to be affected.

Johnson (7, pp. 170-172) illustrates the chi-square transformation by combining two tests of significance (chi square and rho) applied to the same set of data. This example appears questionable because a large correlation coefficient would usually be associated with a large chi square, although the reverse would not necessarily follow.

While the assumption of independence provides the most serious limitation for testing hypotheses concerning combined results, there are additional restrictions which should be noted. For example, while the models may be applied to any complete set of independent results, they obviously are of little value unless the separate tests of significance can be classified together on some rational basis. The application of either model to several experiments with no common aspect would be meaningless. The common problem need not be related to the content of the experimental issue; one might wish to test a hypothesis that a source of bias toward significant results resides in a given method, a particular experimental condition, or a specific laboratory. In the typical instance, however, it is nonsense to combine results bearing upon unrelated experimental problems.

It should be obvious that when applying either model all obtained results must be included in the set

analyzed. The experimenter cannot legitimately select several "promising" or "nearly significant" findings with the hope that together their combined probability will be significant.

We must recognize, of course, that the rejection of the null hypothesis for a *set* of results does not imply rejection of a null hypothesis for each result individually. Thus, if we apply the binomial model and reject the hypothesis that four results, with  $p \leq .05$ , would occur by chance from 25 independent studies, we may not conclude that each of the four is attributable to extra-chance factors. In combining tests of significance, we are testing a hypothesis about the set as a whole, not about the individual members of the set.

Psychologists have sometimes asked: if I have performed two independent experiments and have obtained values at the .10 and .40 levels, why should I not compute the joint probability of obtaining this pair of results by multiplying the two probabilities ( $.10 \times .40 = .04$ )? The difficulty resides in the evaluation of the joint probability. In the interpretation of an individual probability value,  $p$ , we make use of the knowledge that  $p$  has a rectangular sampling distribution, where every value between zero and one has an equal chance of appearing. Such is not the case for products of independent probabilities; it is the  $\chi^2$  transformation which allows assessment of a sampling distribution associated with joint probability.

The error arising from the interpretation of a joint probability as a level of significance can be illustrated intuitively. Consider four independent replications of an experiment, yielding  $p$  values of .90, .60, .30, and .10, under tests of a certain hypothesis.

Clearly these results provide little evidence for rejection of the hypothesis. (The  $\chi^2$  transformation, appropriate to this situation, yields a composite  $p$  greater than .40.) Yet, by finding the product of these four values one obtains  $.90 \times .60 \times .30 \times .10 = .0162$ , which, if erroneously interpreted as a level of significance associated with acceptance of the hypothesis, leads to a completely faulty conclusion.

In more general terms, a statistical test of significance involves the following steps: the selection in advance of an appropriate statistical model for which assumptions are tenable and the sampling distribution of the test statistic is known, the comparison of obtained data (or the statistic derived therefrom) with the model, and the resulting decision that the model does or does not fit the data, i.e., the acceptance or rejection of the statistical hypothesis. It is to be emphasized that the experimenter should specify his statistical model before he looks at his data.

#### THE PROBLEM OF NONINDEPENDENT DATA

Let us consider the problem of evaluating the significance of a set of experimental results when these results have been derived in such a manner as to violate the assumption of independence. When the correlations among the results depart from zero, but are unknown, there is no way in which  $p$  values associated with the individual results may be combined meaningfully. In this case we may, however, achieve an over-all test of significance by treating the set of results as a whole; i.e., we can determine one level of significance for the combined data instead of obtaining  $p$  values for each of the separate results and then combining them.

The suggested approach to this problem is that provided by Hotelling's generalized Student test (6). The most useful form of this test for psychological research probably is that which provides comparison of the sets of means of two samples on each of  $k$  variables. For example, two samples might be selected on the basis of a criterion variable, a sample of  $N_1$  individuals high on the criterion and a sample of  $N_2$  individuals low on the criterion. We have  $k$  variates, with unknown standard deviations and intercorrelations, each of which is measured for the  $N_1 + N_2$  individuals. Hotelling's statistic,  $T^2$ , provides a decision concerning whether the two samples are discriminated significantly on the basis of  $k$  mean differences. Assuming the  $k$  variates to be distributed as the multivariate normal distribution, then the statistic

$$F = \frac{N_1 + N_2 - k - 1}{k(N_1 + N_2 - 2)} T^2$$

is distributed as the familiar  $F$  distribution, with  $k$  and  $N_1 + N_2 - k - 1$  degrees of freedom. The significance of  $T^2$  may be determined readily from tabulated values of  $F$ . While the procedures for finding the value of  $T^2$  are discussed in several texts (e.g., 8, 14), perhaps the most readable account remains the original source (6).

Let us consider the case where  $k$  sets of measurements can be con-

sidered replications of the same variate. When systematic differences among replications can be ignored, the problem of combining separate findings disappears: the  $k$  scores for each individual can be combined and the resulting composite scores can be analyzed by the conventional  $t$  or  $F$  test. When interest also resides in replication effects, one may utilize analysis of variance to test differences not only among groups, but also among replications (9; 4, pp. 288-297). In addition to the assumption that the  $k$  measurements are replications of the same variate, these approaches are further restricted by assumptions of normality, homogeneity of error variance, and independence of variance estimates. The striking advantage of the analysis of variance design, when its use is appropriate, is the relative ease of application when compared with the computational complexity of Hotelling's multivariate procedure.

#### SUMMARY

This paper presents models of two designs by which an experimenter may determine the probability of obtaining a particular set of results on two or more tests of significance. The appropriate assumptions and statistical techniques are discussed. Substitute procedures are suggested for the case when the assumption of statistical independence of the several results is untenable.

#### REFERENCES

1. BAKER, P. C. Combining tests of significance in cross-validation. *Educ. psychol. Measmt.*, 1952, 12, 300-306.
2. BROZEK, J., & TIEDE, K. Reliable and questionable significance in a series of statistical tests. *Psychol. Bull.*, 1952, 49, 339-341.
3. DIXON, W. J., & MASSEY, F. J., JR. *Introduction to statistical analysis*. New York: McGraw-Hill, 1951.
4. EDWARDS, A. L. *Experimental design in psychological research*. New York: Rinehart, 1950.
5. GORDON, M. H., LOVELAND, E. H., & CURETON, E. E. An extended table of chi-square for two degrees of freedom,

for use in combining probabilities from independent samples. *Psychometrika*, 1952, **17**, 311-316.

6. HOTELLING, H. The generalization of Student's ratio. *Ann. math. Statist.*, 1931, **2**, 360-378.
7. JOHNSON, P. O. *Statistical methods in research*. New York: Prentice-Hall, 1949.
8. KENDALL, M. G. *The advanced theory of statistics*. Vol. II. London: Charles Griffin, 1948.
9. KOGAN, L. S. Analysis of variance—repeated measurements. *Psychol. Bull.*, 1948, **45**, 131-143.
10. LINDQUIST, E. F. *Statistical analysis in educational research*. Boston: Houghton Mifflin, 1940.
11. PEARSON, E. S. The probability integral transformation for testing goodness of fit and combining independent tests of significance. *Biometrika*, 1938, **30**, 134-148.
12. PEARSON, K. On the method of determining whether a sample of size  $n$  supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika*, 1933, **25**, 379-410.
13. WILKINSON, B. A statistical consideration in psychological research. *Psychol. Bull.*, 1951, **48**, 156-158.
14. WILKS, S. S. *Mathematical statistics*. Princeton: Princeton Univer. Press, 1943.
15. U. S. DEPARTMENT OF COMMERCE, NATIONAL BUREAU OF STANDARDS. *Tables of the binomial probability distribution*. Applied mathematics series No. 6. Washington: U. S. Govt. Printing Off., 1949.

Received October 27, 1952.

## HISTORICAL NOTE ON THE RATING SCALE

DOUGLAS G. ELLSON AND ELIZABETH COX ELLSON

*Indiana University*

Sir Francis Galton (1822-1911) is generally given credit for the invention of the rating scale or at least for its introduction as a psychological measuring device. Garrett writes, "He [Francis Galton] established one of the first laboratories (in 1884) wherein mental and physical tests—mostly of the sensory-motor sort—could be taken for a small fee; . . . he introduced the rating scale and questionnaire methods later so widely used" (2, p. 58).

Garrett and Schneck say "Historically, the rating scale goes back to Francis Galton, who in 1883 published a scale for rating the clearness of one's mental imagery" (3, p. 103). Guilford also agrees with these statements. In *Psychometric Methods* he writes, "There seems to be little doubt that the first rating scale employed in a psychological problem was that of Galton, used in the evaluation of the vividness of images" (4, p. 264).

We recently found clear evidence that psychological rating scales more highly developed than Galton's were in use when Galton was a child. On a visit to the museum at New Harmony, Indiana, we found a fire-warped bronze or brass plate, accompanied in the display case by a card saying that this was one of the most interesting objects seen by the Duke of Saxe-Weimar on his visit to the New Harmony Colony. As the following material shows, this plate was actually a set of psychological rating scales which was apparently used for a kind of personality evaluation in 1826 and possibly sooner.

The Duke describes this plate in

some detail in his journal, which was published in 1828, *Travels Through North America During the Years 1825-26* (1). In the part of this account describing his visit to the New Harmony Colony, under the date of April 18, 1826, he writes:

Mr. Owen showed me two interesting objects of his invention; one of these consisted of cubes of different sizes, representing the different classes of the British population in the year 1811. . . . The other was a plate, according to which, as Mr. Owen asserted, each child could be shown his capabilities, and upon which after a mature self-examination, he can himself discover what progress he has made. The plate has this superscription: scale of human faculties and qualities at birth. It has ten scales with the following titles: from left to right, self-attachment; affections; judgment; imagination; memory; reflection; perception; excitability; courage; strength. Each scale is divided into one hundred parts, which are marked from five to five. A slide that can be moved up or down shows the measure of the qualities therein specified each one possesses or believes himself to possess (1, p. 121).

The "Mr. Owen" referred to is Robert Owen, who founded the New Harmony Colony in 1825 and left it in charge of his son, Robert Dale Owen, about a year later. Robert Owen had considerable interest in what would now be known as progressive education and child psychology.

Except for a few details, the description of the plate is quite accurate. Its dimensions are approximately  $7 \times 12 \times \frac{1}{4}$  inches. Down the left-hand seven-inch side (which the Duke describes as the top) are listed the traits he mentions, except for a

few (strength, courage, self-attachment) in the upper left corner which are illegible because of fire damage. To the right of each of the ten trait names is a brass strip engraved with 100 scale markings numbered by tens. Before the damage, each strip could be slid outward independently, past an engraved index near the right-hand margin of the plate.

The notion of a rating scale is not the only one anticipated by the plate. If the slides corresponding to the sev-

eral traits were simultaneously placed in positions representing the numerical ratings, they would produce what might now be called a "personality profile." Whether this idea occurred to the inventors or users of the device, we do not know. So far, with the exception of the quotation above, we have found no mention of the plate or of the rating-scale notion in the works of Robert Owen or others who had contact with him.

#### REFERENCES

1. BERNHARDT, K. *Travels through North America during the years 1625-28.* Philadelphia: Carey Lea and Carey, 1828.
2. GARRETT, H. E. *Great experiments in psychology.* New York: Century, 1930.
3. GARRETT, H. E., & SCHNECK, M. R. *Psy-* chological tests, methods, and results. New York: Harper, 1933.
4. GUILFORD, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

Received September 12, 1952.

THE PSYCHOLOGICAL BULLETIN  
Vol. 50, No. 5, 1953

#### A BRIEF NOTE ON ONE-TAILED TESTS

C. J. BURKE  
Indiana University

Concurrent with the recent discussions of one-tailed and two-tailed tests by Hick (4), Jones (5), and Marks (8) there has been a disturbing increase in the use of one-tailed tests in student experimental reports as well as in published and not-yet-published manuscripts. While the popularity of one-tailed tests is undoubtedly attributable in part to the overwillingness of psychologists as a group to make use of the statistical recommendations they have most recently read, there seems to be a certain residual of bad logic, so far as both statistics and psychology are concerned, which merits examination. The writer takes the position already taken by Hick (4) in all important essentials but the argument to be presented differs, at least in emphasis, from that of Hick. It should

be noted that some tests,  $\chi^2$  and  $F$  for example, are naturally single-ended. Nothing said here should be construed so as to apply to them.

Both Jones (5) and Marks (8) seem to the writer to confuse somewhat two quite different notions—that an experimental hypothesis is often directional and that an experimenter may be willing to accept a deviation of any size in the unexpected direction as consonant with the null hypothesis. We shall consider two quotations from Jones.

The model above, the test of the null hypothesis against two-sided alternatives, is the one used most often by investigators in psychology. Yet in many cases . . . it is not the test most appropriate for their experimental problems. More often than not, in psychological research, our hypotheses have a

*directional* character. . . . theoretical considerations allow the postulation of the direction of experimental effects. The appropriate experimental test is one which takes this into account, a test of the null hypothesis against a one-sided alternative (5, p. 44).

It is a fact that many hypotheses in psychological research, experimentally conceived, are directional for the investigator conducting the experiment, but it does not follow from this that one-sided tests should be used in experimental reports.

To amplify these considerations we point out that there are, in many experiments, two statistical decisions to be made and two different levels of confidence may be involved. The first is the decision made by the individual experimenter who frequently plans one experiment from his evaluation of a previous one. We concede that here a one-tailed test is often proper. The second is the decision which determines the place of his findings in the literature of psychology. Here the one-tailed test seems inadmissible. It is the second type of decision with which we are concerned. Marks (8) has in essence repeated from statistical sources a discussion of the Type I and Type II errors which shows that the decisions made in any statistical interpretation depend only upon the underlying populations and the rule of procedure used. Any comparison of alternative rules of procedure must take into account errors of both types, the error of rejecting a hypothesis when it is true and the error of failing to reject it when it is false, but the underlying statistical considerations do not provide automatically a criterion for the selection of one rule over another. Such a criterion is to be sought in the number and kinds of errors the experimenter will tolerate. Roughly, an acceptable criterion

is to make the over-all number of errors as small as possible and at the same time to render large and serious errors relatively impossible. Within the class of hypotheses which are considered to be directional it is likely that a one-tailed test might yield a smaller over-all number of errors than a two-tailed test, but there is, under the single-tailed rule, no safeguard whatsoever against occasional large and serious errors when the difference is in the unexpected direction. If one is less willing to commit a large error than to commit a small one, it does not follow from the theory of testing statistical hypotheses that the experimenter's expectation of a given direction for the result necessarily makes the one-tailed test desirable.

To advance this point in our case against the use of the one-tailed test in the public report, we next take up the second quotation from Jones.

It might be noted that with this formulation of the one-tailed test there is no allowance for the possibility that the true difference . . . is negative. In the type of problem for which the one-tailed test is suited, such a negative mean difference is no more interesting than a zero difference (5, p. 45).

This statement is perfectly correct.<sup>1</sup> If we consider it carefully we discover its import to be that the investigator should use a one-tailed test when he is willing to accept a difference in the unexpected direction, *no matter how large*, as consonant with the hypothesis of zero difference. This is quite a different matter from using a one-tailed test whenever the direction of the difference is predicted, on some grounds or other, in advance. It is to be doubted whether experimental psychology, in its present state, can

<sup>1</sup> In his subsequent discussion, Jones spoils the force of this point by confusing hypotheses to be tested with classes of hypotheses to be guarded against as alternatives.

afford such lofty indifference toward experimental surprises.

The questions raised by the one-tailed test are to be answered finally by considering the effect of general use of this procedure on the content of psychological literature. The writer cannot agree with Hick (4) that its use makes little difference since there is no practical rule for deciding what "significance" really is. In some super-scientific world this point might be well taken, but there is evidence that in our workaday world (where we sometimes read only the concluding sections of reports) it does make a difference whether the investigator has stated that his results were significant. The controversy over the Blodgett effect is a case in point (1, 2, 6, 7, 9, 10, 11, 12).

Remembering that the problem of testing a statistical hypothesis is a statistical problem in which each individual experiment is viewed only as a member of a class of similar experiments and recalling that the properties of any statistical test are determined solely by the procedure followed and by the populations underlying the class, it is pertinent to inquire into the effects of widespread adoption of one-tailed tests upon the literature. The writer believes the following statements to be reasonable forecasts.

1. The discovery of new psychological phenomena will be hindered. Our literature abounds with instances in which the outcome of a given experiment has differed reliably and sharply from expectation. These experiments are usually of great interest—new psychological concepts arise from them. Our science is not yet so mature that these can be expected to occur infrequently. The most recent instance, known to the writer, of conflicting results from experiments thought to be highly similar was re-

ported by Underwood (13) at the 1952 meetings of the American Psychological Association. From any careful examination of contemporary psychological literature we must conclude that nowhere in the field can we have sufficient *a priori* confidence in the outcome of any genuinely new experiment to justify the neglect of differences in the unexpected direction.

2. There will be an increase in barren controversy. Fruitless controversies arise from unreliable results. Conclusions at low levels of confidence tend to be unreliable, and the adoption of one-tailed tests is equivalent to a general lowering of levels of confidence. At a time of severe journal overload this is especially pernicious. There is no substitute in statistical methodology for the carefully designed and controlled experiment in which any important difference between groups will show up at a high enough level of confidence to insure a certain reliability in the conclusion.

3. Abuses will be rampant. It is no criticism of the position held on statistical grounds by Jones and Marks to point out that the considerations involved in the choice of a one-tailed test are really rather delicate. A nice instance of what can happen is seen in an experimental report by Gwinn (3). Gwinn reports two experiments which are not markedly different from each other. They turn out in opposite directions, and, by appropriate selection of the position of his "critical tail," Gwinn establishes significance and near significance (1 per cent and 8 per cent levels, approximately) for his results on the basis of one-tailed tests.

The moral can be pointed with advice. We counsel anyone who contemplates a one-tailed test to ask of himself (before the data are gathered):

"If my results are in the wrong direction and significant at the one-billionth of 1 per cent level, can I publicly defend the proposition that this is evidence of no difference?" If

the answer is affirmative we shall not impugn his accuracy in choosing a one-tailed test. We may, however, question his scientific wisdom.

## REFERENCES

1. BLODGETT, H. C. The effect of the introduction of a reward upon maze performance of rats. *Univer. Calif. Publ. Psychol.*, 1929, **4**, 113-134.
2. BLODGETT, H. C. Reynolds' repetition of Blodgett's experiment on "latent" learning. *J. exp. Psychol.*, 1946, **36**, 184-186.
3. GWINN, G. T. Resistance to extinction of learned fear drives. *J. exp. Psychol.*, 1951, **42**, 6-12.
4. HICK, W. E. A note on one-tailed and two-tailed tests. *Psychol. Rev.*, 1952, **59**, 316-318.
5. JONES, L. V. Tests of hypotheses: one-sided vs. two-sided alternatives. *Psychol. Bull.*, 1952, **49**, 43-46.
6. KENDLER, H. H. Some comments on Thistlethwaite's perception of latent learning. *Psychol. Bull.*, 1952, **49**, 47-51.
7. MALZMAN, I. The Blodgett and Haney types of latent learning experiment: reply to Thistlethwaite. *Psychol. Bull.*, 1952, **49**, 52-60.
8. MARKS, M. R. Two kinds of experiment distinguished in terms of statistical operations. *Psychol. Rev.*, 1951, **58**, 179-184.
9. MEEHL, P. E., & MACCORQUODALE, K. A failure to find the Blodgett effect and some secondary observations on drive conditioning. *J. comp. physiol. Psychol.*, 1951, **44**, 178-183.
10. REYNOLDS, B. A repetition of the Blodgett experiment on "latent" learning. *J. exp. Psychol.*, 1945, **35**, 504-516.
11. THISTLETHWAITE, D. A critical review of latent learning and related experiments. *Psychol. Bull.*, 1951, **48**, 97-129.
12. THISTLETHWAITE, D. Reply to Kendler and Maltzman. *Psychol. Bull.*, 1952, **49**, 61-71.
13. UNDERWOOD, B. J. The learning and retention of serial nonsense lists as a function of distributed practice and intralist similarity. Paper read at Amer. Psychol. Ass., Washington, D. C., September, 1952.

Received October 5, 1952.

THE PSYCHOLOGICAL BULLETIN  
Vol. 50, No. 5, 1953

A NOTE ON THE RECOGNITION AND INTERPRETATION  
OF COMPOSITE FACTORS

WAYNE S. ZIMMERMAN  
*Brandeis University*

French's recently published monograph, *A Description of Aptitude and Achievement Tests in Terms of Rotated Factors* (1), fills a very great need; so great, in fact, that we should expect wide use to be made of it, not only by trained factor analysts, but by counselors, teachers, and personnel workers in business and industry. Particularly because of this anticipated wide use, I was very much disturbed to note the frequency

with which test loadings, reported on a single factor, were actually earned on composite factors. Consequently, as reported, these loadings are in many cases greatly inflated. The serious distortion that results colors the monograph throughout, but is especially apparent in Part II.

In Part II, brief descriptions of factors are presented, accompanied by test names listed in the order of magnitude of their loadings. This

section will more than likely receive the greatest application, especially by those interested solely in selecting the best test of a particular factor. It is probably safe to assume that few of the relatively untrained, and perhaps not all of the trained, people will check carefully to see if the particular test concerned achieved its loading in analyses where a full complement of factors in the related sub-domains was isolated.

In his introduction French discusses composite factors, but in such a way that the reader may be left with the impression that these factors occur very infrequently, or that when they do their nature is apparent and they do not therefore seriously influence the monograph presentation.

Composite factors may occur when several tests included in a correlational matrix have a similar factor pattern. If, at the same time, no other tests have a sufficiently disparate pattern on the same factors, a separation between factors cannot be achieved, even when an ample number of centroid factors are extracted.

Composite factors may be produced also when a curtailed number of centroid factors are extracted. In too many instances tests have been included for analysis which actually represent a greater number of factors than can be extracted according to any criterion applied to determine when the extractions should cease. In many other instances a too rigid criterion is employed. The experience of some analysts has shown that the use of additional centroids, extracted after the most liberal statistical criterion is satisfied, frequently brings about more readily a psychologically meaningful solution.

In this short critical note, I can do no more than call attention to one or two of the most glaring examples of the failure to note the accretion be-

cause of composite factors to certain reported test loadings.

On the Space factor it is necessary to work down the list through eleven tests before locating one with a loading that seems to be relatively free of the composite taint. The loading for the leading test is based upon a single analysis only—an analysis in which just four factors were extracted with only one in the spatial-visualization domain. AAF analyses of a nearly similar test (1, AFL-7; 2, No. 32) have demonstrated that where other related factors (i.e., Perceptual Speed, Length Estimation, Visualization) are isolated, the test variance is spread among them leaving a very much reduced weight on the Spatial factor. The position of the next three tests in the list can be questioned by reference to this same list where a combination of the three tests into a single form is ranked twenty-fourth, with considerably less than one-half of the variance. This circumstance suggests that there is probably a significant amount of specific variance represented when any of the three highly related tests are inter-correlated and analyzed as single tests. At the same time a much reduced Space loading was found for this combined form in analyses where both Visualization and another factor in the spatial-visualization area (1, Factor SO) appeared also (1, AFM-21, AFN-21). The next seven tests, which include several of the form-board type, were all analyzed in batteries in which only a single spatial-visualization factor was isolated. Where more than one related factor was revealed, the Space variance of similar tests was reduced very markedly. (See 1, ThA-21, Fr-10, M, B-14. Also AFE-15, AF6-8, AFM-23.)

Of the first ten variables in the list of Deduction tests, five fairly obviously are composites, since three

represent grades assigned in academic subjects and two are based upon comprehension of school subject matter. Two other tests actually achieved the loadings credited to them on a factor that was called Attention by the author of the analysis. One test was analyzed with two other forms of the same test, both of which produced one-half or less variance on the same factor. The remaining two test loadings were based on a single score derived on three different tests, each of which singly in other analyses demonstrated a different and complex factor pattern. In none of the analyses in which these first ten tests were analyzed was a representative group of factors in the reasoning area isolated.

It seems that there are still factor analysts who, although they accept rotation as a necessary step in the process, somehow are not aware of one implication from the pattern (often indistinct and sometimes enmeshed beyond recognition, it must be granted) that tends to form in the centroid loadings. It is well known that in dealing with intellectual variables the first centroid loadings are usually all positive and the second centroid, as well as those that follow, divide positive and negative variance equally. What is apparently overlooked is the tendency for the second centroid to split the most obvious dichotomy, the third centroid to split the next most obvious dichotomy, and so on. For example, if the

battery contains both linguistic and quantitative tests, the second centroid will most likely separate these two major groups. In order to effect further separations between test variables on the linguistic end of this second axis (which, in pair-by-pair rotation, has now been extended by rotation with Centroid Axis I) it must be rotated with the next centroid that shows this new separation most clearly. Similarly, further separation of the quantitative tests must be achieved through rotation of the quantitative end of the second axis (also extended through its rotation with Centroid Axis I) with a centroid which in turn shows the most distinct separation among the quantitative tests. This separating process must be continued until all the factors are accounted for. If too few centroid factors are extracted, some of these separations cannot be made. Consequently, a test which actually contains variance on two or more factors may appear with all of that variance confined on a single factor.

It is my feeling that the failure to give composite factors the attention they merit must be considered either a serious oversight or a serious error of omission. Since the user must make his own evaluation, I hope this note will draw attention to the necessity of considering very carefully the serious effect of composite factors upon certain test loadings reported in French's monograph.

## REFERENCES

1. FRENCH, J. W. *The description of aptitude and achievement tests in terms of rotated factors*. Chicago: Univer. of Chicago Press, 1951.
2. GUILFORD, J. P., FRUCHTER, B., & ZIMMERMAN, W. S. Factor analysis of the

Army Air Forces Sheppard Field battery of experimental aptitude tests. *Psychometrika*, 1952, 17, 45-68.

Received August 14, 1952.

## A REJOINDER TO ZIMMERMAN'S NOTE

JOHN W. FRENCH  
*Educational Testing Service*

Zimmerman's criticism of the treatment of composite factors in *A Description of Aptitude and Achievement Tests in Terms of Rotated Factors*<sup>1</sup> brings up an important point. Analyses using very few tests often yield combination factors that seem to be composed of two or more factors ordinarily found separately. The tests found to have high loadings on these combination factors are often complex. When such complex tests are analyzed in larger batteries, they often receive lower loadings on each of the separate factors.

Now, what should be done about this in comparing loadings from many analyses, large and small? The centroid of the high-loading tests for a particular factor in one analysis is probably never exactly duplicated in any other analysis. Nevertheless, it is necessary to make some cross-identifications, provided the factors are close enough to being alike. How close is that? If you insist upon extreme closeness, you identify so few factors that the result is not useful. If you are very liberal, the factor loadings cannot be compared because they are not loadings on the same thing. The combination factors mentioned by Zimmerman cause most of the trouble, but factors do not always combine equally. For example, consider the case of an analysis in which neither of the well-known factors A or B appears, but in which tests ordinarily found on A and B are now found on the same factor. As usually happens, the division is not exactly equal; the A-factor tests may be more numerous and highly loaded. Perhaps a combination factor correlating .90 with A and .40 with B can safely be identified as A, while a combina-

tion factor correlating .80 with A and .60 with B should be left unidentified. Even in the former case a test having loadings on both A and B will, unfortunately, tend to have a higher loading on the impure A factor than do the tests which best represent pure A. Thus, a compromise must be made between too few identifications, on the one hand, and too poorly representative loadings, on the other.

Zimmerman's note suggests that listing the tests according to the size of their loadings is misleading. They were listed in this manner to be entirely objective. How else should they be listed? Perhaps they should have been listed in accordance with some value derived from differentially weighting such considerations as size of loadings, number of tests in the analysis, number of analyses in which the test appears, number of subjects, estimation of closeness of factor to other factors identified by the same name, closeness of the test in psychological meaning to that of the factor, etc. To do this would have been to put more reliance on my judgment with regard to weights than I cared to put, and would have increased the effects of any possible misinterpretations of mine beyond the degree in which they already appear.

One answer, as suggested by Zimmerman, is to wave more red flags. I am now preparing a similar exposition of personality factors, where the problem of combined factors is even worse. I will have to give this problem much attention, and, because of his note, will take extra pains to warn the nonfactorists of possible pitfalls. I wish Zimmerman were right in saying that the monograph serves the masses, but the figures on sales indicate that we have a most exclusive group of readers.

Received October 21, 1952.

<sup>1</sup> FRENCH, J. W. *A description of aptitude and achievement tests in terms of rotated factors*. Chicago: Univer. of Chicago Press, 1951.

## BOOK REVIEWS

DE GRAZIA, SEBASTIAN. *Errors of psychotherapy*. Garden City: Doubleday, 1952. Pp. 288. \$3.00.

LINDNER, ROBERT. *Prescription for rebellion*. New York: Rinehart, 1952. Pp. 305. \$3.50.

These two books, when read together, make an interesting contrast. After a lapse of a week or two, however, they become strangely intermingled in memory.

Each is a sort of intellectual *tour de force*. Each is the work of a literate, passionate prophet who, starting with relatively little data, evolves a logical structure which becomes a formula for saving humanity from itself. Both authors have little good to say for contemporary psychotherapy and each sets forth his convictions concerning the only way things may be set aright before it is too late. The fact that their peregrinations lead to markedly different conclusions does not dissipate the inhibition the memory of one book has on the memory of the other.

Both authors are sharply critical of contemporary psychology and psychiatry. They are critical, though for different reasons, of traditional psychoanalysis and nondirective therapy, and they heap scorn on shock therapy and lobotomy.

Both authors mistake hypotheses for conclusions. Each ends his book with generalizations which, in more modest men, might lend themselves to testable hypotheses. Hoping for data, one runs head-on into the back cover.

de Grazia's message is fairly simple. Neurosis is a moral disorder. The unmoral psychotherapists, worshipping false gods, do not possess the kind of authority essential for forgiveness of moral transgression. For-

giveness is essential, however, to valid therapy—mere toleration degenerates whatever moral standards remain. The book ends with a description of the true psychotherapy which, it turns out, requires periodic confession to a religious authority who has power to recognize and forgive moral transgression. In the process of building his logical structure, de Grazia examines the status of confession in various religions and the consequences of its loss or retention. About the most convincing datum bearing on the value of confession is a quotation from a Hawthorne novel. De Grazia debunks the available data on evaluation of therapy (what is cure?) and he exposes the error in such concepts as unconscious motivation, psychological determinism, developmental factors in neurosis, and other sacred psychological cows. To some psychologists he will not appear to lean over backwards to his friends. In his introduction, he anticipates that some will say that he has set psychotherapy back twenty-five hundred years. This estimate may be high.

de Grazia writes with a sharp stylus. Lindner uses a half-inch camel's-hair brush and mixes more than the usual amount of oil with his paints. Lindner, too, is critical of modern psychotherapy, but for a different reason. To him psychotherapists are succeeding in making people adjusted. This is terrible. Adjustment is conformity and stereotypy and leads to Mass Man and totalitarianism. Lindner shares with his readers his discovery that the goal of true psychotherapy and, more broadly, of social therapy is rebellion. He builds his case for individuality and creativity. He decries passivity and homogeneity. He places psychologists at

the head of the list of the culprits leading mankind into bondage under the banner of adjustment. He exposes psychology's hypocrisy and criticizes its guiding fictions. In the end he reveals his prescription for the new psychotherapy. Psychotherapy must have as its goal and criterion of success "the re-entry of the individual . . . into the evolutionary stream." He does not specify the exact units to be used in measuring this variable.

Many psychologists reading either of these two books will find themselves muttering objections and writing comments in the margins. Despite this they will find much that is stimulating. Both authors express their ideas clearly and with great certainty. Their readers may be less certain but they cannot avoid having their thinking provoked.

GEORGE W. ALBEE.

*Washington, D. C.*

BENEDEK, THERESE: *Studies in psychosomatic medicine: psychosexual functions in women*. New York: Ronald, 1952. Pp. x+435. \$10.00.

This is the second volume in the series of psychosomatic studies sponsored by the Chicago Psychoanalytic Institute. The first eleven chapters reprint the 1942 monograph published by the National Research Council. The last four chapters are psychoanalytical essays on some female psychosexual functions. Since no new data, methodology, or analyses are appended to the reprint, the evaluation of the major portion of this book is in essence a review of the 11-year-old monograph.

Two earlier studies led Benedek, and her collaborator Rubinstein, to conclude that in the sexual cycle there is a dynamic correlation between each hormonal variation and the psychodynamic manifestations of

the sexual drive; that the sexual cycle could be adequately described only if the emotional cycle and the hormonal cycle were both included as integral components.

The present study is an extension of the earlier investigation. Daily psychoanalytic therapy productions and daily physiological hormone indices were collected from 15 psychologically disturbed patients. These data were combined with the data of the two earlier studies making a total of 152 female sexual cycles. Six of the patients contributed two-thirds (101) of the cycles. Although the vaginal smear was not a direct hormone assay, biophysiological studies by Rubinstein gave reasonable assurance that the smear was an adequate first approximation of the hormone level.

The daily therapy data were psychoanalytically interpreted, and classified in terms of three groups of predominant "psychodynamic tendencies." From these psychodynamic tendencies the daily hormonal characteristics were predicted: (a) Estrogen production was inferred when the patient presumably showed active psychodynamic tendencies directed towards an externalized object (e.g., heterosexual tendencies, masculine identification, active homosexual tendency, infantile sexuality, oedipus complex, masochism-sadism, and exhibitionistic tendency). (b) Progesterone hormone levels were inferred when passive receptive psychodynamic tendencies could be interpreted from the daily material (e.g., libidinous narcissism, genital receptivity, retentive wishes, pregnancy wishes, mother-conflict and mother-identification, motherliness, and passive homosexual tendency). (c) Low hormone levels were inferred when the therapy data permitted the interpretation of such psychody-

namic tendencies as destructive hostility, negative narcissism, infantile demanding dependence, depressive and eliminative tendencies.

Certain methodological unclarities suggest caution in accepting the author's claim of high predictive accuracy. Benedek does not state whether the individual days were analyzed in random order, nor how the influence of knowledge about related days in the cycle was guarded against in the evaluation of the given day. These points are particularly pertinent inasmuch as Dr. Benedek herself states that the interpretation of any psychological production requires a knowledge not only of the given sexual cycle, but also of the complete clinical material for the patient. There are no data as to the reliability of the criteria for determining psychodynamic tendencies either for one judge or for independent interpreters. It is also regrettable that no definite statement is made as to the specificity of the hormone predictions which the psychological material could be expected to yield. At times it appears from the data in Tables 17-42 that a three-point scale is used (estrogen, progesterone, low hormone level); yet at other times a more elaborate scale is used (incipient estrogen, estrogen, high estrogen; incipient progesterone, progesterone, declining progesterone, low hormone level). This creates a problem in understanding how the author arrived at her summary statements of accuracy of prediction. The criteria for partial and total discrepancy between predicted hormone level and vaginal smear hormonal index are sometimes difficult to interpret: psychological prediction *Progesterone*, vaginal smear finding *Estrogen* is listed as a partial discrepancy (p. 322); psychological prediction *Estrogen and Minimal Pro-*

*gesterone*, vaginal smear finding *Increasing Estrogen* is listed as no discrepancy (p. 293).

There seems little doubt, however, that Dr. Benedek was able to infer the probable hormonal status from the psychological productions of these particular subjects at far beyond chance levels. This is an important clinical demonstration of psychosomatic unity of function. But the lack of scientific modesty in limiting the broad generalizations and conclusions derived from the data casts a shadow over the positive contribution of this monograph.

The first of the four essays, *Mother-Child as a Primary Unit*, deals with "... the psychodynamics of the symbiosis which exists during pregnancy, is interrupted at Birth, but remains a functioning force, directing and motivating the mental and somatic interaction between mother and child." The hormonal functions which are related to milk secretion induce an emotional attitude in the mother in which active and passive receptive tendencies predominate in a manner comparable to the progesterone phase of the feminine sexual cycle. This is said to promote the development of motherliness and to foster the development of the personality of mother and child by gratification of the dependent needs of both. The apparent regression to an oral level by the mother in her unconscious identification with the infant presumably is the psychodynamic tendency which predisposes to depressive reactions in the puerperium.

The *Climacterium*, topic of the second essay, is characterized as a period of intrapersonal reorganization in women. The adequacy of the earlier psychosexual adjustment from menarche through motherhood becomes

the important source of "surplus gratification." Thus, the end of the propagative function may be viewed as the release of nonsexualized energy for a new phase in which an active, adaptive life plan can be developed.

The third paper, *The Functions of the Sexual Apparatus and Their Disturbances*, is a potpourri of psychoanalytic clichés and a recapitulation of the essential conclusions of the monograph and of the first two essays in this series of four. The discussion of dysmenorrhea is the only new contribution. Dysmenorrhea is treated without differentiation as to whether it is primary or secondary in character. The psychoanalytic explanation of dysmenorrhea is in terms of a diminished control of the ego at the time of menstruation over psychosexual conflicts which by "returning from repression" mobilize anxiety and general nervous system reactions and predisposes the woman to overreact to the premenstrual-menstrual hormone changes.

The final essay, *Some Psychophysiological Problems of Motherhood*, follows the pattern of reviewing the monograph and the first essay in the series. The generalizations about "motherliness" conflicts appear to be derived from upper-middle class women who can afford psychoanalytic treatment, yet these generalizations are freely expanded to include all women. There is a high degree of autonomy in these essays, and their juxtaposition in print seems more an accident of convenience than a plan.

The republication of the 1942 monograph to make it more available to a wide audience is to be welcomed. The clinical methodology, the efforts to elaborate criteria for judging psychodynamic tendencies, and the demonstration of psychosomatic unity of functioning are of value to research

psychologists, physiologists, and physicians. The brief summaries of psychoanalytic theory introduced for the nonpsychiatric reader are excellent condensations, succinct yet thorough.

M. ERIK WRIGHT.  
*University of Kansas.*

MAIER, NORMAN R. F. *Principles of human relations*. New York: Wiley, 1952. Pp. ix + 474. \$6.00.

This text on human relations extends the trend of publications in this field. Industrial personnel now realize more than ever before the need for scientific techniques and approaches in the promotion of better employer-employee relations. Very often, a human relations program in industry is based on the experiences of the personnel man or upon half-thought-out methods provided by "pulp" magazines in the field of supervision and management. Industry has grown tired of these approaches and now wants something scientific upon which to base its programs. Maier's text fulfills this need to some extent.

The book has been aimed primarily at the industrialist. It is typically a Maier book on frustration, morale, and attitude-change techniques. Approximately 90 per cent of the book is devoted to attitude-change techniques. Little attention is placed on experimental results. It is in this area that the greatest weakness is shown. The author tends to draw too many conclusions and overgeneralizes from single case studies. A feeling develops that techniques and approaches are colored by the background and training of the author. Basic studies by Lewin, Bavelas, Coch and French, and Lippitt are cited, but are these enough to pave the way for the application of the group decision methods in industry?

Maier's chapters on role-playing and group-decision procedure are undoubtedly the best written in the field. Possible applications to industrial personnel problems are well made. The book will be well received in industry. He has many case studies to bring home his points.

One wonders why nondirective counseling is discussed in his last chapter. Unfortunately, this introduces the reader to a complicated area of human study that might have best been eliminated or included as one of the interview methods.

In summary, the book is very timely in its presentation and undoubtedly is what is now needed in industry. It will provide personnel people with some concrete basis for this field of study.

D. J. MOFFIE.

*North Carolina State College.*

BROWER, DANIEL, AND ABT, LAWRENCE E. (Eds.) *Progress in clinical psychology*. Vol. I. New York: Grune & Stratton, 1952. Section 1: pp. xi+328; \$5.75. Section 2: pp. xxiii+236; \$5.00.

This is the first volume, in two sections, of a projected series of periodic reviews of the literature in clinical psychology. Section 1 contains Parts I, II, and III: Introduction, Diagnostic and Evaluative Procedures, and Psychotherapy. Section 2 contains Parts IV, V, VI, and VII: Developmental Processes, Applications of Clinical Psychology to Special Areas, Approaches to Clinical Psychology, and Professional Issues. The range of the material encompassed is so great as to preclude a detailed critical evaluation. One can, however, examine the basis for selection of material, the general theoretical approach, and perhaps comment briefly on some of the 42 chapters.

In the preface to the first section, the editors state that "We early reached a decision to employ value judgments in the selection of material to be covered. . . ." Aside from the subsequent statement that "We have included what we believe to be a judicious selection of the varied, complex and interlaced applications of clinical psychology in special areas of concern and interest . . .," the reader is nowhere told what these values are. If the content of this volume is a reflection of the editors' value judgments, a psychoanalytic orientation predominates. This is probably inevitable since psychoanalysis has provided the most valuable hypotheses for psychodagnosis and psychotherapy. However, other points of view are represented, as in the chapter by Raymond B. Cattell on *P* technique and the conventional discussion of intelligence testing in the child by Dale B. Harris, among others.

It is the psychoanalytic orientation apparently which leads Abt, in the opening chapter, to attribute the delay in the development of psychology relative to the other sciences to the fear that men have of discovering their own natures. In the reviewer's opinion, this is a naive view of the history of science, and represents a lack of sobriety in the application of psychoanalytic concepts, which has been abandoned even by many psychoanalysts.

As can be expected in any large collection of papers, they vary widely in quality. There are a number of excellent critical reviews of the literature. For example, the chapter on Psychosurgery by Birch is outstanding because of the conciseness of its exposition and the acuteness of its insight into the basic lack of sound supporting data. With particular reference to the increasing use of psychosur-

gery as a method of choice rather than of last resort, Birch concludes that "the evidence indicates the need to halt the ever-widening use of a radical practice that has neither a clear theoretical justification nor a sound empirical base."

Hertz's chapter on the Rorschach is a thorough and competent review of the literature. There is an absence of defensiveness with respect to experimental data which makes it particularly appealing. At the same time, she points out that the need for experimental research should not blind the clinician to the equal need for *art* and intuition in the interpretation of an individual Rorschach protocol. In sharp contrast stands the rather desperate defense of the Szondi test by Deri, who attacks all critical studies, and omits consideration of some of the most damaging ones. Her plea for tolerance of ambiguity by clinicians thus fails to strike a sympathetic chord.

Clinical and other applied psychologists should find a number of the individual papers informative, particularly those on topics with which they may deal only peripherally in their everyday practice. For this reviewer, the chapters on Measures of Aptitude, Achievement and Interest; Testing for Psychological Deficiency; House-Tree-Person and Human Figure Drawing; Mental Deficiency; and Cultural Anthropology, by Gustav, Hunt, Brown, Heiser, and Sargent, respectively, were in this category. It is doubtful, however, that the general academic psychologist will find much of lasting theoretical value which cannot be obtained elsewhere. Abt's chapter on The Emergence of Clinical Psychology is interesting but sketchy. Mowrer's chapter on Neurosis and Its Treatment as Learning Phenomena is a restatement of his

well-known position on the "neurotic paradox." Else Frenkel-Brunswik's chapter on Social Psychology emphasizes the contributions of psychoanalysis to this area.

Some of the authors cite references and findings without adequate comment. The reader is often left in the midst of a maze of numerous journal references. The two chapters by Ellis on Self-Appraisal Methods and Sexual Disorders are saved from this fate, despite the inclusion of more pages of closely printed references than of loosely printed text, by the author's summaries and integration.

On the whole, there is much information of interest in these two books, most of which, however, can be found in other integrated sources, such as the *Annual Review of Psychology*. The question may be raised therefore whether these volumes add sufficiently to available sources to warrant a new series of periodic reviews of the literature, even if in a specialized area.

JULIUS WISHNER.  
*University of Pennsylvania.*

LANSING, A. I. (Ed.) *Cowdry's problems of ageing*. (3rd Ed.) Baltimore: Williams and Wilkins, 1952. Pp. xxiii+1061. \$15.00.

To the psychologist, the extent to which psychological data are omitted from this book is amazing. While our contributions to the understanding of ageing have not been spectacular, they are not nonexistent as would appear from this volume. No mention is made of researches on age changes in speed and accuracy of response, learning, creativity, memory, vocabulary, or mental abilities of any sort. Of 40 chapters, not one is psychological.

Some fields closely related to psychology are likewise ignored. For ex-

ample, although there are three chapters on the sexual organs, sexual behavior is barely mentioned. I am unable to find any reference to Kinsey. Likewise there is no treatment of psychopathology, although a sharp increase in mental disease in later maturity is one of the prominent features of ageing.

The orientation of the volume is clearly medical; but medicine, it seems, does not include psychiatry. The treatise is not purely biological and medical; there are four chapters by sociologists and economists, but it could scarcely have been the richness of the literature on ageing in these disciplines which prompted the inclusion of these chapters. To illustrate, the chapter by Havighurst on Roles and Status of Older People cites only three investigations. Two of these are by Havighurst and one is a University of Chicago dissertation. Such poor documentation is not, however, characteristic of the medical and biological chapters.

The heterogeneity of this book illustrates the fact that there are no gerontologists, but only men specializing in the gerontological aspects of particular fields. In this situation, it would seem to have been wiser had the decisions concerning areas to be included been made by a board of editors representing many disciplines.

WAYNE DENNIS.

*Brooklyn College.*

PIÉRON, HENRI, PICHOT, PIERRE, FAVERAGE, J. M., AND STOETZEL, JEAN. *Méthodologie psychotechnique.* (Book II in *Traité de psychologie appliquée*.) Paris: Presses Universitaires de France. Pp. viii + 208. 1.000 fr.

This is the second volume of an ambitious treatise on applied psychology, published under the editor-

ship of Henri Piéron and planned in seven volumes. The first volume, issued in 1949 and entitled *Differential Psychology*, dealt with the concepts of individual variability and "types," aptitudes and their factorial analysis, and the role of heredity. The present volume is concerned principally with the methods of applied psychology.

Piéron's section contains a general introduction, concerned both with tests as tools of measurement and with the evaluation of the test scores, and an inventory of principal psychometric tests. The European reader would have profited by being given a reference to sources of information on the vast amount of American work on the construction and utilization of psychological tests, other than Whipple's Manual (1910), Bronner *et al.* (1928), Garrett and Schneck (1933), and Rapaport (1945). References to individual tests and to works on test theory are more up-to-date. Piéron's chapter will acquaint the American student with developments in the field of testing in Europe, especially in France. Of particular interest and potential usefulness is a nonverbal (pictorial) group test of intelligence designed for children of school age and developed by René Gille.<sup>1</sup> The test was standardized on a sample of about 100,000 French boys and girls aged 6 to 12 years and was adapted for use with the African Negro children. Selected tests of sensory, perceptual, motor, and a variety of "mental" functions (such as attention, memory, and spatial orientation) are described, mostly very briefly.

<sup>1</sup> Cf. Heuyer, G., Piéron, H., Piéron, Mme. H., and Sauvy, A. *Le niveau intellectuel des enfants d'âge scolaire* (Intellectual level of children of school age). Paris: Institut National d'Etudes Démographiques, 1950. Rep. No. 13 of *Travaux et documents*.

Methods for the study of personality are presented by Pichot. The methods are classified as (a) analytical (direct observation, whether during an interview or in a partially structured group situation; questionnaires, almost exclusively American; and objective tests of personality), and (b) syncrétique or projective (their rationale and an inventory of methods, following L. K. Frank's system of classification). Frequently, a technique is simply mentioned and the reader is referred to the original publication for description. The French reader will seek more detailed information in the author's recent book on mental tests in psychiatry (*Les tests mentaux en psychiatrie: I. Instruments et méthodes*. Paris: Presses Universitaires de France, 1949).

Stoetzel devotes 40-odd pages to methods for the study of opinion, with references limited almost entirely to the American literature. He notes that in French psychological literature the appearance of the very word "opinion" is very recent, missing in the large *Traité de Psychologie* edited by Georges Dumas and used in Lalande's comprehensive *Vocabulaire de la Philosophie* (1932) only in a non-technical meaning. Of interest is the analysis of the reasons for a delay in the cultivation of this field, especially in France: Cartesian tradition for which only "clear and distinct" ideas merit the scientist's consideration; biological orientation of scientific psychology; lack of appreciation for the role of social interaction; disregard of value judgments; domination by the theoretical models of physics; and concern for function rather than content. Stoetzel also brings up the point that institutional changes are likely to result in anxiety to which the threatened groups will respond aggressively, treating the *study* of the

changes in the same way as they react to the prospect or presence of the changes themselves. The three chapters of the section deal with methods (attitudes of an individual, opinions of a group), some results obtained by the application of these methods, and the historical development of opinion research. For a more detailed treatment one should consult the author's other publications (*Théorie des Opinions*, 1943; *L'Etude Expérimentale des Opinions*, 1943; *Les Sondages d'Opinion Publique*, 1948).

Statistical methods used in applied psychology were described by Faverge. The chapter is brief (pp. 259-302), especially in reference to its ambitious content covering, as it does, both analysis of variance and factor analysis. The presentation is limited to concepts and formulae. Computational procedures are not given in a workable detail.

The potential user of the treatise would do well to arm himself with the *Vocabulaire de la Psychologie* (dictionary of psychology), published by the Presses Universitaires de France in 1951. Otherwise, he will look in vain in standard dictionaries for an English equivalent or the meaning of numerous terms (such as "docimologie," the scientific study of scholastic and other examinations). The heavy impact of the English language on the terminology of French scientific psychology—with some expressions taken over literally (tapping, closure, the term *test* itself), others in translation (*échantillonage aréolaire* for "area sampling," *sondages d'opinion* for "public opinion polls," *ensemble* for "population")—is likely to make an occasional use of the dictionary mandatory for the French student as well.

JOSEF BROŽEK.

University of Minnesota.

GILBERT, JEANNE G. *Understanding old age*. New York: Ronald Press, 1952. \$5.00.

In today's society, physicians, psychologists, nurses, and social workers are called upon to solve the practical problems presented by older people and their relatives. Each of these professions has acquired a vocabulary and a body of information peculiar to itself. Furthermore, biologists, biochemists, physiologists, psychologists, and sociologists who have studied problems of aging have described their results in the technical terminology and within the concepts of their own scientific discipline. It is high time that workers from each profession concerned with aged people be offered a summary of our present knowledge of aging as gleaned from all scientific disciplines written in intelligible language. The preparation of such a volume is a formidable assignment to which Dr. Gilbert has addressed herself.

The first part of the book is devoted to a cataloging of physical, physiological, endocrine, emotional, intellectual, and social-relation changes with increasing age. Findings are presented in serial fashion with little selection for significance in terms of behavior. The reader is offered little help in assessing the importance or meanings of the changes reported. For instance, the occurrence of senile plaques in the cerebral cortex of individuals suffering from senile dementia prior to death is discussed (p. 24), but the work of Malamud showing that there is no relationship between the incidence of senile plaques in the brain and impairment of mental function is not mentioned. The reader is not informed (until p. 293) that in many cases of senile dementia, the structural changes in the brain are not

sufficient to account for the psychosis. Considerable emphasis is placed on the importance of diet in age changes, but no mention is made of the important work of McCay on the relationship between dietary restriction and longevity. In the discussion on heredity, no reference is made to the important work of Raymond Pearl. Although the Third (1952) Edition of *Cowdry's Problems of Ageing* was probably not available at the time the manuscript was written, it is surprising to find that all references to Cowdry are to the First (1939), rather than to the more complete Second Edition (1942). Neither will the professional worker be particularly impressed by the inclusion of Ripley (Believe It or Not) as a bibliographical reference. The treatment of technical terms is not consistent, since many of the simpler ones are defined (thrombosis, lumen, epithelium), whereas others are not (colloid of the thyroid gland; "hold" and "don't hold" tests). The section on intellectual changes seems preoccupied with the concept of deterioration and fails to emphasize the findings that decline is minimal or even absent in intellectual functions that are practiced and used in the personal life of the aging individual. In discussing age changes in creative ability, the author leans heavily on anecdotal material rather than on the careful studies of Lehman and his co-workers.

Part two of the book deals with abnormal life changes in aging, but no clear definition of abnormal is made. Actually this section deals with clinically diagnosable disease states found in elderly people. Physicians will no doubt look askance at the definite statements of symptoms and methods of treatment of diseased states. A good many errors of fact

have crept into the presentation. Although there is a great deal of theorizing about emotional factors and their causal relations to disease states, the author leads the unwary astray in presenting these theories in the same way as factual observations.

The third section of the book, concerned with professional work with the aging, contains a great deal of common sense presented in a readable fashion. The section is specific and offers many valuable suggestions on the theme that old people can be integrated into effective community life. The blind acceptance of the value of vitamin supplements is a limitation. Of special value is emphasis on the idea that although we as yet have no panaceas for reversing the physical changes of aging, much can be done on the emotional and personal reactions of individuals toward these changes.

NATHAN W. SHOCK.

*Baltimore City Hospitals.*

REIK, THEODOR. *The secret self.* New York: Farrar, Straus and Young, 1952. \$3.50.

Theodor Reik here continues the process, elaborated in *Listening with the Third Ear*, of being both the psychoanalyst and his couch. In the present volume he gathers some of his early ruminations on diverse topics and expands upon them in free-associational contemplation. The points of departure range from lamb chops on a restaurant menu and thoughts on Goethe during an analytic session, to hypnagogic interpretations of parent-child relationships within Reik's family.

The psychoanalyst may look with favor on the richness of commentary developed around each leitmotiv, but the scientifically oriented reader will look askance at the methodology and its rationale. Juggling phrases, al-

lusions, and emotional content, Reik attempts to explore the unconscious motives of the heroes, authors, or prime-movers in each of his vignettes. Unfortunately, the phrases, allusions, and emotional content are not those of the heroes and authors, *but of Reik himself* as he responds associatively to clues which perhaps he alone perceives. What justifies this approach? To quote Reik (e.g., page 170), "These impressions are elusive and allusive, too, but the analytic technique will, I hope, help us in attributing a new significance to them, and in evaluating them in the service of psychological understanding."

The reviewer was tempted by this justification to interpret in similar fashion the handy and obviously Freudian slip occurring in the chapter on *Sexual Symbolization in Modern Plays*. On page 225, Reik declares (or was it the typesetter?): "It is possible to speak of 'emotional uticaria' or puritis." Freely associating, the reviewer perceives an error common to both medical terms used —the omission of the "r" in each. The connotation is immediately clear—it is reminiscent of those days when he (and obviously Reik, too) was frustrated by the lack of edible oysters in the devastating "r"-less months of the year. The culturally determined sexual symbolization is seen to be thematically related to Brocchi's *Traité de l'Ostréiculture*. . . .

The book evokes—this gentle scoffing notwithstanding—a genuine admiration of the cultural erudition which Reik brings to his analyses, an erudition applied with practiced familiarity and without ostentation. It is with regret that one must, in the face of this, accuse Reik of a major modern crime—guilt by free association.

DAN L. ADLER.  
*San Francisco State College.*

## *Outstanding* **McGRAW-HILL Books**

### **DEVELOPMENTAL PSYCHOLOGY**

By ELIZABETH B. HURLOCK, University of Pennsylvania. *McGraw-Hill Publications in Psychology*. 556 pages, \$6.00

With important stress upon each level of development as foundation for the next, this text covers the life span from conception to death with emphasis on outstanding characteristics in each major life period. Close correlation between mental and physical growth and methods of change in interests, attitudes, and behavior are discussed. Included also is a review of major experimental studies.

### **PERSONALITY AND ADJUSTMENT**

By WILLIAM L. PATTY, Los Angeles City College, and LOUISE S. JOHNSON, University College, Rutgers. 403 pages, \$4.75.

Combining detailed discussions of the scientific approach and practical application to various areas of living, this fine text offers a simple, practical introduction to personality adjustment and mental hygiene. Research from the fields of psychology, psychiatry, physiology, sociology, mental hygiene, and education is applied to the understanding of the individual.

*Send for copies on approval*

**McGRAW-HILL BOOK COMPANY, INC.**

330 West 42nd Street

New York 36, N.Y.

---

## OUTSTANDINGLY SUCCESSFUL TEXTS IN PSYCHOLOGY

---

### *Psychology of Personal and Social Adjustment*

Henry Clay Lindgren, San Francisco State College

"It is well-written and shows keen perception of students' interest. The combination of emphasis on understanding personality development and on 'practical' psychology is very useful."—John Pierce-Jones, University of Oregon

"Sound, highly readable, informative text, with an intelligent clinical orientation."—Hirsch L. Silverman, Rutgers University

"It's good to see a psychology text for 'normal people'."—William P. Matthews, Lynchburg College

### *Social Psychology • An Interdisciplinary Approach*

Hubert Bonner, Ohio Wesleyan University

"This is a text we have been waiting for. We shall adopt it this coming fall. Comprehensive, scholarly, and readable." Peter P. Klassen, University of Illinois

### *A Laboratory Manual for Social Psychology*

Wilbert S. Ray, Trinity College, Hartford, Conn. • INSTRUCTOR'S GUIDE

"... carefully wrought, and expressed in crisp and clear English. The field research studies, added to the laboratory experiments, take the student into the practical aspects of research."—Christian O. Weber, Wells College

### *Psychology*

Henry E. Garrett, Columbia University

"This is an excellent introductory book. It reads well, but it is not so over-simplified as to lose the essential meaning of the principles involved. I particularly like the liberal use of diagrams, charts, tables, and the like."—James H. Mailby, Southern Methodist University

### *Psychology in Use • A Textbook in Applied Psychology*

• SECOND EDITION

Edited by J. Stanley Gray, University of Georgia

"There is absolutely no substitute in the field of applied psychology for this integrated, correlated presentation by one who knows his field and has had much experience in teaching it."—Richard Trumbull, Syracuse University

## American Book Company

COLLEGE DIVISION

Cincinnati • Chicago • Boston • Dallas • Atlanta • San Francisco

55 FIFTH AVENUE, NEW YORK 3, NEW YORK